

GenCore version 5.1.4 ps 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 27, 2003, 21:11:47 / Search time 8407 Seconds

(without alignments)  
1942.033 Million cell updates/sec

Title: US-09-821-687-4

Perfect score: 2982

Sequence: 1 MATHEVNGNGTEEPMDTSTA.....AQQGRGGKGVEAGPDLIQ 561

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-Q/cgcn2.1/USPTO\_spool/US09821687/runat.22042003.165241.7951/app.query.fasta\_1.711  
-DB=GenEmbl -QFMT=fastap -SUFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -List=45  
-DOCALLIGN=200 -THR SCORE=DOT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09821687 @CGCN 1.1 5173 @runat.22042003.165241.7951 -NCPU=6 -ICPU=3  
-WARN -NO -MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb\_da:\*  
2: gb\_hvg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sv:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_scs:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_hvg\_hum:\*  
31: em\_hvg\_inv:\*  
32: em\_hvg\_other:\*  
33: em\_hvg\_mus:\*  
34: em\_hvg\_pln:\*  
35: em\_hvg\_rdt:\*  
36: em\_hvg\_mam:\*  
37: em\_hvg\_vrt:\*  
38: em\_by:\*  
39: em\_hvg\_hum:\*  
40: em\_hvg\_mus:\*  
41: em\_hvg\_other:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2982	100.0	3452	10 AB035725	AB035725 Mus muscu
2	2964.5	99.4	2221	9 AF155568	AF155568 Homo sapi
3	2964	99.4	1686	9 AY034483	AY034483 Homo sapi
4	2922	98.0	1872	9 AY034481	AY034481 Homo sapi
5	2922	98.0	2932	9 AF037448	AF037448 Homo sapi
6	2883	96.7	2373	10 AF093821	AF093821 Mus muscu
7	2882	96.6	2079	6 AR177851	AR177851 Sequence
8	2752	92.3	3319	9 BC032643	BC032643 Homo sapi
9	2705.5	90.7	1767	9 AY034482	AY034482 Homo sapi
10	2676	88.7	103819	9 HS0775C13	AL109618 Human DNA
11	2676	89.7	212659	2 AL390737	AL390737 Homo sapi
12	2604	87.3	2232	10 AF408434	AF408434 Mus muscu
13	2502	83.9	3101	6 AX011753	AX011753 Sequence
14	2413.5	80.9	1899	10 AF441128	AF441128 Mus muscu
15	2403	80.6	2644	9 AP000364	AP000364 Homo sapi
16	2391.5	80.2	2722	9 BC001449	BC001449 Homo sapi
17	2275	76.3	2471	10 BC004679	BC004679 Mus muscu
18	2114	70.9	181810	2 AC126148	AC126148 Rattus no
19	2066.5	69.3	121304	2 AC079134	AC079134 Mus muscu
20	2066.5	69.3	196488	2 AC084410	AC084410 Mus muscu
21	1745	58.5	139986	9 AL161799	AL161799 Human DNA
22	1520	51.0	221309	2 AC122217	AC122217 Mus muscu
23	1468	49.2	151297	2 AC118904	AC118904 Rattus no
24	1382	46.3	934	6 AX013715	AX013715 Sequence
25	1360	45.6	129237	2 AC127932	AC127932 Rattus no
26	1301	43.6	2606	3 AY058477	AY058477 Drosophila
27	1175	39.4	202374	10 AC122868	AC122868 Mus muscu
28	1172.5	39.3	110783	2 AC129693	AC129693 Rattus no
29	1159.5	38.9	118429	2 AC119124	AC119124 Rattus no
30	1041	34.9	165311	2 AC112849	AC112849 Rattus no
31	964	32.3	2473	9 AF262323	AF262323 Homo sapi
32	963.5	32.3	2498	10 BC013481	BC013481 Mus muscu
33	959	32.2	2218	4 CFA38855	AJ388555 Canis fam
34	955.5	32.0	2585	10 BC034195	BC034195 Mus muscu
35	908.5	30.5	4020	9 AK000280	AK000280 Homo sapi
36	903	30.3	159699	2 AC026833	AC026833 Homo sapi
37	902.5	30.3	188295	9 AC098869	AC098869 Homo sapi
38	902.5	30.3	248201	2 AC115293	AC115293 Mus muscu
39	896	30.0	1863	10 AF290984	AF290984 Rattus no
40	894	30.0	1785	10 AY028945	AY028945 Rattus no
41	883.5	29.6	1794	9 HSA27078	AJ27078 Homo sapi
42	881.5	29.6	1945	9 AF209192	AF209192 Homo sapi
43	879.5	29.5	1770	9 HSA27079	AJ27079 Homo sapi
44	876.5	29.4	1761	9 AF271789	AF271789 Homo sapi
45	876.5	29.4	128739	2 AC116602	AC116602 Mus muscu

RESULT 1

#### ALIGNMENTS

AB035725  
LOCUS AB035725 3452 bp mRNA linear ROD 31-MAR-2000  
DEFINITION Mus musculus SYNCRIP mRNA, complete cds.  
ACCESSION AB035725  
VERSION AB035725.1 GI:6576814  
KEYWORDS SYNCRIP.  
SOURCE Mus musculus (strain:DDY) postnatal 21 day cDNA to mRNA.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS Mizutani, A., Fukuda, M., Ibata, K., Shiraiishi, Y., and Mikoshiba, K.  
TITLE SYNCRIP, a cytoplasmic counterpart of heterogeneous nuclear  
ribonucleoprotein R, interacts with ubiquitous synaptoctagmin  
isoforms  
J Biol. Chem. 275 (13), 9823-9831 (2000)  
JOURNAL  
MEDLINE 20200483  
REFERENCE 2 (bases 1 to 3452)  
AUTHORS Mizutani, A., Fukuda, M. and Mikoshiba, K.  
TITLE Direct Submission  
JOURNAL Submitted (11-DEC-1999) Akihiro Mizutani, Brain Science Institute,  
RIKEN, the Laboratory for Developmental Neurobiology, 2-1 Hirosawa,  
Wako, Saitama 351-0198, Japan (E-mail: amizutani@ims.u-tokyo.ac.jp,  
Tel:+81-48-467-9745, Fax:+81-48-467-9744)  
FEATURES  
source location/Qualifiers  
1..3452  
/organism="Mus musculus"  
/strain="DDY"  
/db\_xref="taxon:10090"  
/dev\_stage="postnatal 21 day"  
1..3452  
/gene="SYNCRIP"  
1..3452  
/gene="SYNCRIP"  
/evidence=experimental  
154..1839  
/gene="SYNCRIP"  
/protein\_start=1  
/protein\_id="BA088342.1"  
/db\_xref="GI:6576814"  
/translation="MATEHVNNGTETPEMDTSAVHSENFQTLDAQLPQVKAETLD  
EIVVAGLVNASHSDDERAIEALKEFNEDGALVLAQFDSPLSHVOKSAFLCGVMKTY  
RORRKGSTVADSSKDEDAKILALERTYTLDTTGKRGSGPPDSVSGQPSV  
GTEIFVKGKIPDLFEBELVPLFRKAGVIMPLRLMDPLTGANKGVAFTVPTCEAAOB  
AVLVNNEHRSKHIQVCI SVANNLIFVSI PKSTKQILIEBPGKVBGLPDVILY  
HOPDDKKRGRGPFLEIEBKTKAQAARRRLMSGKVKYGVNGVTEWADPIEDDPVW  
AKQVLFVRNLAVTVEELIEKSPQGLERVKLKDVAFHFDERDGVKAMEBN  
GKDLBENIEIVPAKPPDQKREKAKQRAKQMDYDYVGPVPHNPPTRRGRGG  
RGGYYPDYGYEDYGYDYNHNRGGYEDPYGYEDFOVGARGGRGARGAAP  
SRGRGAAPRGARGYSGORGGPGSARGVARGAGAAQOQRGKGVARGPDLQ"

BASE COUNT 1056 a 623 c 777 g 996 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3,36e-131 Length: 3452  
Score: 2982.00 Matches: 561  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-821-687-4 (1-561) x AB035725 (1-3452)

QY 1 MetAlaThrGluHisValAsnGlyAsnGlyThrGluGluProMetAspThrThiSerAla 20  
DB 154 ATGGCTCAGAAACATGTTAAAGGAAATGCTACTGAGAGCCCACTGATCACTTCCGCA 213

QY 21 ValIleHisSerGluAsnPhenGlnThrIleuAspAlaGlyLeuProGlnIleValAla 40  
DB 214 GTTATCATTCAGAAATTTTCAGACATGCTGCTGCTGCTTTTACCAACAGAAAGTGGCT 273

QY 11 GlnIleuLeuAspGlnIleuValAlaGlyLeuValAlaHisSerAspLeuAspGluArg 60

DB 274 GAAAACTGATGAAATTTACGTTCAGAGGCGTACTTGCACATGATGATTTAGTAGAAG 333  
QY 61 AlilIleGluAlaLeuIleuGluPheAsnGluAspGlyAlaLeuAlaValIleuGlnIlePhe 80  
DB 334 GCTATCAGAACCTTTAAAGAGTTCAATGAAAGCGGCGCATTTGGCAGTGCCTTCAACGTTT 393  
QY 81 LysAspSerAspLeuSerHisValGlnAsnIleSerAlaPheLeuGlyValIleMetLys 100  
DB 394 AAGACAGATATCTCTCTCAATGTCAGAACAAAGAGCTTTTATGTGAGTCATGAG 453  
QY 101 ThrTyArgGlnArgGluIleuGlnGlyThrIleValAlaAspSerSerIleuGlyProAsp 120  
DB 454 ACTTACAGCAGAGAGAAAGAAACAGGGGACCAAGTACAGACTCTGTAAGAGCAGAT 513  
QY 121 GlnAlaIleuIleuGlnAlaLeuLeuGluIleuGlnIleuGlyThrIleuAspValThrThrGly 140  
DB 514 GAGCAAGATTAATTAAGGCACTTTTGAAAGAACAGGCTTACACACTTATGATGATCAAGGT 573  
QY 141 GlnArgLysThrGlyIleuGlyProProAspSerValTySerGlyIleGlnIleProSerVal 160  
DB 574 CAGAGAGATGATGAGAGACCACTCCAGATCCGTTTATTCAGGTACAGACCTTCTGTT 633  
QY 161 GlyThrGluIlePheValGlyIleProArgAspLeuPheGluAspGluLeuValPro 180  
DB 634 GGCACTGAGATATTTGTGGGAAAGATCCCAAGATCTGTTTGAAGATGAGCTTGTCCA 693  
QY 181 LeuPheGluIleuValAlaGlyProIleTPAspLeuArgLeuMetLysProIleuThrGly 200  
DB 694 TTTATTTGAGAAACCTGACCTATATGATGATCTTCTGTTATGATGATGATCGCTCAGTGT 753  
QY 201 LeuAsnArgGlyTyraIlePheValThrPheCysThrIleGlnAlaIleGlnIleVal 220  
DB 754 CTCAACAGAGTTATGAGTGTTCCTGCTTGTCTTGTACAAAGAACACACAGAGGCTGTT 813  
QY 221 LysLeuTyraAsnAsnHisGluIleArgSerGlyLysHisIleGlyValCysIleSerVal 240  
DB 814 AATCTGATATATATCATGAAATTCGTTCCGGGAAACACATGGTGTCTGCACTCAGTT 873  
QY 241 AlaAsnAsnArgLeuPheValGlySerIleProLysSerIleThrIleGluGlnIleLeu 260  
DB 874 GCACAACATATGCTTTTGTGGGCTCGATTCCTCAAGAGTAAACCAAGGACGATTCCTT 933  
QY 261 GlnGluPheSerIleValThrGluGlyLeuThrAspValIleLeuTyraHisGlnProAsp 280  
DB 934 GAGGAATTTGCAAGATGACAGAGGCTTCACAGATCTCATTTTATACCAACCACTGAT 993  
QY 281 AspLysLysLysAsnArgGlyPheCysPheLeuGluTyrgluAspHisLysThrAlaAla 300  
DB 994 GACAAAGAAAAAACAAGAGCTTTGCTTCTTGAATATGAAATCATCAAAAACAGCTGCC 1053  
QY 301 GlnAlaArgArgArgLeuMetSerGlyLysValIleValIlePheValGlyIleThrVal 320  
DB 1054 CAGGCAAGACGTAGCTAATGAGTGTGAAAGTCAAGTCTGGGGAATGTGGAACCTGAT 1113  
QY 321 GlnTyraIleAspProIleGluAspProAspProGluValMetAlaIleValIleValLeu 340  
DB 1114 GAGTGGCTGATCTTATTGAAGTCTGATCTGAACTTATGCAAGGTAAGTAAGTGC 1173  
QY 341 PheValArgAsnLeuAlaAsnThrValThrGluGluIleLeuGluIleuLysSerPheSerGln 360  
DB 1174 TTTGTACGCAACCTTGCACAAACGTTACAGAAAGAAATTTTGAAGAAAGCATTTATTCAG 1233  
QY 361 PheGlyLysLeuGluIleuValIleLysIleuLysLeuLysAspTyraIlePheIleHisPheAspGlu 380  
DB 1234 TTTGGGAAATCGAAGAGAGTAAGAAAGCTAAAGATTAATGCTTTCATTCATTTTGTATGAG 1293  
QY 381 ArgAspGlyValAlaValIleValIleMetGluIleuMetAsnGlyLysAspLeuGluIleuAsn 400  
DB 1294 AGAGATGGTGTCTCAAGGCTATGAGAAAGTAATGAATGATTAAGCTTGGAGGAGAAAT 1353  
QY 401 IleGluIleValPheAlaLysProProAspGlnIleuArgGlyIleuArgLysAlaGlnArg 420  
DB 1354 ATTGAATGTTTCTTGTAAAGCAACAGATCAGAAAGGAAAGAAAGAAAGCTCAGAGG 1413

QY 421 GlnAlaAlaLysAsnGlnMetTyrAspAspTyrTyrTyrTyrTyrProPheHisMetPro 440  
 DB 1414 CAGAGCAAGAAAGATCAAAAGTATGATGATTAATCACTATTATGCTCACTCATAGCCT 1473  
 QY 441 ProPheThrArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGly 460  
 DB 1474 CCCCACCAAGAGAGGTGAGGGGCTGAGAGGTAGAGGTGAGGTAGATGATGATGATGAT 1533  
 QY 461 TTTGTTTGTGTAAspTyrTyrAspTyrTyrAspTyrTyrAspTyrTyrAspTyr 480  
 DB 1534 TATGATATCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1593  
 QY 481 TTTGTTTGTGTAAspTyrTyrAspTyrTyrAspTyrTyrAspTyrTyrAspTyr 500  
 DB 1594 TATGATATCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1653  
 QY 501 ArgGlyAlaArgGlyAlaAlaProSerArgGlyArgGlyAlaAlaProProArgGlyArg 520  
 DB 1654 AGAGGAGCAAGGGGCTGCTGCTCATCAGAGGTGAGGGGCTGCTGCTGCTGCTGCTG 1713  
 QY 521 AlaGlyTyrSerGlnArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGly 540  
 DB 1714 GCGGTTATTCACAGAGAGAGGCGCTGATCAGCAAGAGCGCTGCGGTTGAGAGAG 1773  
 QY 541 GAlaAlaGlnGlnArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArg 560  
 DB 1774 GGTGCCCAACAAAGAGGCGCGGGGAGAAAGGGGTGAGGCGGCTGCTGCTGCTG 1833  
 QY 561 Gln 561  
 DB 1834 CAA 1836

RESULT 2  
 AF155568 2221 bp mRNA linear PRI 10-JUN-1999  
 LOCUS Homo sapiens NSAP1 protein (NSAP1) mRNA, complete cds.  
 DEFINITION AF155568  
 ACCESSION AF155568  
 VERSION AF155568.1 GI:5031511  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 2221)  
 HARRIS, C.E., BODEN, R.A. and ASCELLO, C.R.  
 A novel heterogeneous nuclear ribonucleoprotein-like protein interacts with NS1 of the minute virus of mice  
 J. Virol. 73 (1), 72-80 (1999)  
 JOURNAL 9847309  
 MEDLINE 99102562  
 PUBMED 9847309  
 REFERENCE 2 (bases 1 to 2221)  
 HARRIS, C.E., BODEN, R.A. and ASCELLO, C.R.  
 Direct Submision  
 Submitted (01-JUN-1999) Biochemistry and Molecular Biology, UBC, 2146 Health Sciences Mall, Vancouver, BC V6T 1Z3, Canada  
 JOURNAL Location/Qualifiers  
 TITLE 1..2221  
 FEATURES  
 source  
 1..2221  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /note="cell line"  
 1..2221  
 /gene="NSAP1"  
 205..1893  
 /gene="NSAP1"  
 /note="similar to Homo sapiens hn RNP R; probable RNA binding protein; interacts with NS1 of the minute virus of mice"  
 /codon\_start=1  
 /product="NSAP1 protein"  
 /protein\_id="AAD38198.1"  
 /db\_xref="GI:5031512"  
 /translation="MATEHVNGNGTEBPMDTTSVAVHSENFQTLIDAGLPQVAKELD

BASE COUNT 699 a 402 c 561 g 559 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,41e-130 Length: 2221  
 Score: 2964.50 Matches: 559  
 Percent Similarity: 99.82% Conservative: 2  
 Best Local Similarity: 99.47% Mismatches: 0  
 Query Match: 99.41% Indels: 1  
 DB: 9 Gaps: 1  
 US-09-821-687-4 (1-561) x AF155568 (1-2221)  
 QY 1 MetAlaThrGluHisValAsnGlyThrGluGluProMetAspThrSerAla 20  
 DB 205 ATGGCTACAGAAATGTTATGGAATGGTACTGAAGGCCATGGATGATCTTCTGCA 264  
 QY 21 ValIleHisSerGluAsnPheGlnThrLeuAspAlaGlyLeuProGlnLysValAla 40  
 DB 265 GTTATTCATTCAGAAATTTTCAGACATTCGTTGATGCTGCTTACACAGAAAGTGGCT 324  
 QY 41 GluLysLeuAspGluIleTyrValAlaGlyLeuValAlaHisSerAspLeuAspGluArg 60  
 DB 325 GAAATCTGATGATTAATTTTACCTTGAGGCTGAGTGGACATGATGATTAATGAAAGA 384  
 QY 61 AlaIleGluAlaLeuLysGluPheAsnGluAspGlyAlaLeuAlaValLeuGlnInphe 80  
 DB 385 GGTATTTGAAGCTTTAAAGATTAATGAAAGCGGTGATGGCAGTTCTTCAACAGTTT 444  
 QY 81 LysAspSerAspLeuSerHisValGlnAsnLysSerAlaPheLeuGlyValMetLys 100  
 DB 445 AAGACAGATCTCTCTCATGTTCAAGAACAAAGTCTTTTATGAGTCAATGAG 504  
 QY 101 ThrTyrArgGlnArgGluLysGlnGlyThrLysValAlaAspSerSerLysGlyProAsp 120  
 DB 505 ACTTACAGCAGAGAGAAACAGAGACCAAGTACGATTCAGTGAAGACACAGAT 564  
 QY 121 GlnAlaLysIleLysAlaLeuLeuGlnArgThrGlyTyrThrLeuAspValThrThrGly 140  
 DB 565 GAGGCAAAATTAAGGCACTCTTGAAAGAACAGGCTACACCTTGATGTGACCTGGA 624  
 QY 141 GlnArgLysTyrGlyLysProProAspSerValTyrSerGlyGlnGlnProSerVal 160  
 DB 625 CAGAGAGATGAGAGACCACTCAGATTCGTTTATTCAGTCAAGACCTTCTGTT 684  
 QY 161 GlyThrGlnIlePheValGlyLysIleProArgAspLeuPheGluAspGluLeuValPro 180  
 DB 685 GGACCTGAGATATTTGTGGAAAGATCCCAAGAGATCTATTTGAGAGATGAACCTTGTCCA 744  
 QY 181 LeuPheGluLysAlaGlyProIleTyrAspLeuAspLeuMetMetAspProLeuThrGly 200  
 DB 745 TTTATTTGAAAGCTGAGACCTATGAGATCTCTCTATATGATGATGATGATGATGAT 804  
 QY 201 LeuAsnArgGlyTyrAlaPheValThrPheCysThrLysGluAlaGlnGlnLysVal 220  
 DB 805 CTCATATAGAGGTTATGCGTTTGTCACTTTTGTACAAAGAGACAGCTCAGAGGCTGTT 864  
 QY 221 LysLeuTyrAsnAsnHisGlnLysSerGlyLysHisIleGlyValCysIleSerVal 240  
 DB 865 AACTGTATTAATCAAGAAATGCTTGTGAAACATATTTGGTGTGCAATTCAGT 924  
 QY 241 AlaAsnArgLeuPheValGlySerIleProLysSerLysThrLysGlnGlnIleLeu 260  
 DB 925 GCACCAATATAGGCTTTTGTGGCTCTATCTTATGAGATGAACAGAGACAGATTCTT 984

QY	261	985	QY	281	Db	1045	QY	301	Db	1105	QY	321	Db	1165	QY	341	Db	1225	QY	361	Db	1285	QY	381	Db	1345	QY	401	Db	1405	QY	421	Db	1465	QY	441	Db	1525	QY	461	Db	1585	QY	481	Db	1645	QY	501	Db	1705	QY	521	Db	1765	QY	541	Db	1825	QY	560	Db	1885	RESULT 3	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORD	SOURCE
	Gluc1pHspSerLeuValThrGluGlyLeuThrAspValIleuLeuYrHisGlnProAsp	GAAGAATTATAGCAACTAACAGAGGCTCTTACAGAGTCACTTTATACCAACACCGGAT		ApplVbLysAlaSerAsnArgGlyPheCySPheLeuGluTyrGluAspHisLysThrAlaAla		GACCAAGAAAAAACAAGAGGCTTTTGTCTTCTGTAATTAAGAATCAAAAACAGCTGCC		GlnAlaArgArgArgLeuMetSerGlyLysValLysValTProGlyAsnPalGlyThrVal		CAGGACAGGCGTAGGTTAATGAGTGTAAAGTCACAGCTCGGGGAGATGTGGAACTGTT		GlutPAlaAspProIleGluAspProAspProGluValMetAlaLysValLysValLeu		GAATGGCGCTGATCCTTATAGAAGTCTGTATCTCGAGGTTATGGCAAAAGTAAAGCTG		PheValArgAsnLeuAlaSerThrValThrGluGlnIleLeuGlnLysSerPheSerGln		TTTGGGAACTGGACAGAGTGAAGAAAGTTAAAGATTATAGCTTCACTCATTTTGATGAG		PheGlyLysLeuGlnArgValLysLysLysLeuLysAspTyrAlaPheIleHisPheAspGlu		TTTGGGAACTGGACAGAGTGAAGAAAGTTAAAGATTATAGCTTCACTCATTTTGATGAG		ArgAspGlyValAlaValLysAlaMetGluGluMetLeuAsnGlyLysAspLeuGluGluAsn		CGAGATGGTGCTGTCTCAAGGCTATGAGAAAGATCAATGCGAAAGACTTGGAGGGAGAAAT		IleGluIleValAlaPheAlaLysProProAspGlnLysArgLysGluArgLysAlaGlnArg		ATTGAATATGTTTGTGGCAAGCACCAAGATCAAGAAAAAGAAAGAAAGAAAGCTCAGAG		GlnAlaAlaLysAsnGlnMetTyrAspAspTyrTyrTyrTyrGlyProProHisMetPro		CACGACGCAAAAATCAATGATGATGACGATTACTACTATTATAGTTCACCTCATATGCC		ProProThrArgGlyArgGlyArgGlyGlyArgGlyGlyTyrGlyTyrProProAspTyr		CCTCCAAACAAGGCTCGAGGGGCTGAGAGGTAGAGGCTTATAGATATCTCCACAGTTAT		TyrGlyTyrGluAspTyrTyrAspTyrTyrGlyTyrAspTyrHisAsnTyrArgGlyGly		TATGATATGAAGATTAT		TyrGluAspProTyrTyrGlyTyrGluAspPheGlnValAlaGlyAlaArgGlyArgGlyGly		TATGATATGATCATTAT		ArgGlyAlaArgGlyAlaAlaLysProSerArgGlyArgGlyAlaAlaLysProProArgGlyArg		AGAGAGACGAAAGGGGTCTCTCCATCCATCAAGAGGTCGTGGGGCTCTCTCCCCGGGTACA		AlaGlyTyrSerGluAspGlyGlyLysProGlySerAlaArgLysValAlaArgGlyAlaArgGly		GCGGATATATCAACAAGAGAGAGTCTCTGATCAGCAAGAGCGCTTCAGAGTGTGAGAGGA		GlyAlaGlnGlnGlnArgGlyArgGly--GlyLysGlyValAlaGlnAlaGlyProAspLeu		GGTGCCCAACAACAAGAGGCGCGGGGCAAGGAGAAAGGGGTGAGGCGCGTCTGACTGTG		LeuGln		TTACAA	AY034483	AY034483	Homo sapiens hnRNP Q1 mRNA, complete cds.	AY034483	AY034483.1	GI:15809589	Homo sapiens.

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 1686)
AUTHORS	Moutrelatos,Z., Abel,L., Yong,J., Kataoka,N. and Dreyfuss,G. SNM interacts with a novel family of hnRNP and spliceosomal proteins
JOURNAL	EMBO J. 20 (19), 5443-5452 (2001)
MEDLINE	21458434
PUBMED	11574476
REFERENCE	2 (bases 1 to 1686) Moutrelatos,Z.
AUTHORS	Direct Submission
JOURNAL	Submitted (11-MAY-2001) HMI/ Biochemistry, University of Pennsylvania, 330 CRB, 415 Curie Blvd., Philadelphia, PA 19104, USA
FEATURES	Location/Qualifiers
SOURCE	1..1686 /organism="Homo sapiens" /db_xref="taxon:9606"
CDS	1..1686 /codon_start=1 /product="hnRNP Q1" /protein_id="AAK59705.1" /db_xref="GI:158095930" /translation="WATHEVNGNTESEPMDTTSVAIHSENPQTLLDAGLPQKAEKID ELIVYAVASHSDUDERAIETALKENFEGALAVIQEFDSLSHVQNKSAPLCGVWKTY RORFOGCTKVADSKSPDEAKIKALERTYTTLDTTGQRKGPPDPDSVYSQQPSV GTPEVFKEIPRDLFEDELVPLEFACPIIMLRMMDDLITGLNGYAFVFCYEAAGE AVLKYNNHRSRGKHIGVCISVANNRIFVSGIPIKSKOEOLIEFSFVTEGLTDVLILY HODDPKKNRSFGCLEYEDHKHTAOARMSRKVYMGVNGVATEMDPLEDDPRTVM AKRYVLFVRNLANTVTREILEKAPSQSGKERKKLDYAFIRHDERDGAVKAMEENK KGDIENGINETIVFAKPDPDQKRERKAKOAKOMDYDYTPGPMPPTRRGRGGCG RGSGEPDPDYGYEDVDYGYGYHYHNRRQGEEDPYGEDPFVGARRGGRGARGAAP SRHGGAAPPGRAGYSQSGPGSGARGVGARGAQAQQRGGRGVGEAPDLIQ"
BASE COUNT	536 a     282 c     435 g     433 t
ALIGNMENT SCORES:	
Pred. No.:	1.12e-130     Length: 1686
Score:	2964.00     Matches: 1588
Percent Similarity:	99.64%     Conservative: 1
Best Local Similarity:	99.47%     Mismatches: 2
Query Match:	99.40%     Indels: 0
DB:	Gaps: 0

US-09-821-687-4 (1-561) x AY034483 (1-1686)

Oy	1	MetAlaThrGluHisValaSnGlYAsnGlYThrGluGluPProMetaSerThrThrsEra	1	20
Dd	1	ATGGGTACAGAACAATTGTAATGAAAATGCTACTGAAGGCCCATGAGACTACTTCCTGCA	60	
Oy	21	VallLeHisSerGluAenPhcGlnThreuleuAapalagilyLeuProGlnLyvaAla	40	
Dd	61	GTTATCCATTACAGAAATTTTCAGACATTCCTTGATGCTGGTTAACCAAGAAAGTGCT	120	
Oy	41	GluysLeuaSpGluileTyerValAlaGlyyeuvAlaahisseraSpLeuaSpGlaArg	60	
Dd	121	GAATAACTAGTAGAATTTCCTGTCAGGGCTTAGTGCACATAGTATTGATGAAGA	180	
Oy	61	AlailegualaleuLyseGluPhaasnGluApsGlyalaleuaAlavalleuGlnGlnphe	80	
Dd	181	GCTATGTAAGCTTTAAAGATTCATGAAGAACGCGTGATGGCAGTCTTCAACAGTTT	240	
Oy	81	LysaasPeraApLeuSerHisvalaGlnaenLysseraAlaPhelEuCyBglyValmeLys	100	
Dd	241	AAAGCACTGATCTCTCATGATTCAGAAACAAAGTGCTTTTATGTGAGATCATGAAG	300	
Oy	101	ThrTyraTgaGlnatvgGluLyegInglYthrTyerValAlaaesPeraSerLyegIProasp	120	
Dd	301	ACTTRACGGCAGAGAGAAAAACAAGGACCMAAGTAGAGATTCTAGTAAAGACACAGAT	360	
Oy	121	GluAlaLySilleyalaLeuGluLuarGrhIslYTrThreulaSpValThThrlslY	140	

QY	501	ArgGlyAlaArgGlyAlaAlaProSerArgGlyArgGlyAlaAlaProProArgGlyArg	520
Db	1501	AGAGAGCAAGGGGGCTGCTTCATCCAGAGAGTGTGGGGCTGCTCTCCCCGGGTRGA	1560
QY	521	AlaGlyTyrSerGlnArgGlyGlyProGlySerAlaArgGlyValArgGlyAlaArgGly	540
Db	1561	GCCGGTATTATCAAGAGAGAGGTCTTGATCAGCAAGAGCGTTCCAGGTGCCAGAGGA	1620
QY	541	GlyAlaGlnGlnGlnArgGlyValArgGlyGlyValGlyAlaGlyAlaGlyProAlaGln	560
Db	1621	GGTGCCCAACAACAAGAGCGCCGGGGTAAAGGGGTCAAGCGCGGTCTGACTGTTA	1680
QY	561	Gln 561	
Db	1681	CNA 1683	
RESULT 4			
LOCUS	AY034481	1872 bp	RNA
DEFINITION	Homo sapiens hnRNP Q3 mRNA, complete cds.	linear	PRI 03-OCT-2001
ACCESSION	AY034481		
VERSION	AY034481.1	GI:15809585	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteleia; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1872)		
TITLE	Mourelatos,Z., Abel,L., Yong,J., Kataoka,N. and Dreyfuss,G.		
JOURNAL	SNM interacts with a novel family of hnRNP and spliceosomal proteins		
MEDLINE	EMBO J. 20 (19), 5443-5452 (2001)		
PUBMED	21458434		
REFERENCE	2 (bases 1 to 1872)		
AUTHORS	Mourelatos,Z.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-MAY-2001) HML/ Biochemistry, University of Pennsylvania, 330 CRB, 415 Curie Blvd., Philadelphia, PA 19104, USA		
FEATURES			
Source	Location/Qualifiers		
	1..1872		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	1..1872		
	/product="hnRNP Q3"		
	/protein_id="AKS59703.1"		
	/db_xref="GI:15809586"		
	/translation="MATEHNGNGTEBPMDTTSAVIHSENFOTLLDAGLPQKAEKLD EIVYALVANSIDLDERAIEALKKFNEDGALAVIQQRKDSLSHVQNSAFLLCGWMTY RQREKGTAVADSSKGPDEAKTALRLERTGYTLDTTGQRKYGAPPDSVYSGQPSV GVLTPVGGKIPRLDEFLVPLPEFKAPLINDLRIMDBPLTELNGVAFTVCTKEAOE AGLIYNNHIEIRSKGHIQVCISSVNNRLFYGSIIRKSTKEQILIEFSKTEGLTDVLY HODDDKKKNSPCFLEVEDHKTAAQARRRLMSKRYWGVGEVEMADPIEGDDPEVM AKRVKLFVRLANTVETIELIEKAFSQGLERKYLKDYAFIHEDRDGAVKMEEN GRLDEENIEIVAKPPDQKREKKAQQAQAKQMTDDTYTTPPHMPPTRRKGKG RGRTGCPEDTYTGEDTYDYGIDYHNYRGGEDPYTGEDFVGAGRGAGRGAGGAP SRGAGTAPKRGAGYSGQRGPGSARGVRGAGYQOORGGVGAGRGAGRGAGGGRK ADEYNPDPSRRRTNNQNMGSQPIAQQPLQGGDHSGNYGYSNORFYDTFQOWK"		
BASE COUNT	590 a 323 c 489 g 470 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.15e-128	Length:	1872
Score:	2922.00	Matches:	549
Percent Similarity:	99.10%	Conservative:	2
Best Local Similarity:	98.74%	Mismatches:	5
Query Match:	97.99%	Indels:	0
DB:	9	Gaps:	0



```

/codon_start=1
/product="Grx-rbp"
/protein_id="AACT2926.1"
/db_xref="GI:3037013"
/translation="MATEHYNGNGTEPEMDTTSVAVHSNFTQLDAGLPQVYAEKLD
EIVVAQLVASHSDDERALKEPNEDGALAVIQPKSDLSHVQKSAFLGVMKTY
RQREKGTIVADSSKSGDEAKI KALLERGTLDVTYGRKYGPPDPDVGQOPV
GHEIFVWKIPDLFEDELVPLEKAGPIWDLRLMDPLTGLNRGYAFVTEKAAOE
AVLYNNHHSRSGKIHGVCSIVANNRLFVGSIPKSTKEGIIIEEPKTEGJLDVILY
HOPDKKRSRPFLEVEDHKTAQAARRLSMKRYVMNVCSTENADPIEDPDEVM
AKYKULPYRNLANATTEETLEKAFSOFGLERVKIKDYAFIHPDRDGAVAMEEMN
GKDLBSENEIIVFAKPPDGRKREKAQRAANQMDYDYGYPPMPPTKRGK
RGQYGPDPYDYGEDYDYGYDYNHRYRGYEDPYGYGAFVGAAGRGGRAGAP
SRGRGAPPRGRAGSQRGPGSARVGORGRGVRGARAGRGAGVGGKREK
ADGYNOPDSKRQTNQNMWSOPIAQOPFGGDSHSNVYKSENGEFYQDTRGQOMK"
BASE COUNT      868 a      520 c      823 g      721 t
ORIGIN
Alignment Scores:
Pred. No.:      1,82e-128      Length:      2932
Score:          2922.00      Matches:      549
Percent Similarity: 99.10%      Conservative: 2
Best Local Similarity: 98.74%      Mismatches: 5
Query Match:     97.99%      Indels:      0
DB:              Gaps:      0
US-09-821-687-4 (1-561) x AF037448 (1-2932)
QY      1 MetAlaThrGluHisValaIaenGlyAsnGlyThrGluGluProMetAspThrThrSerAla 20
DB      526 ATGGCTACAGAAACATGTATATGAAATGTACAGAGAGCCATGATACCTTCGCA 585
QY      21 ValIleHisSerGluAsnPhenGlnThrLeuLeuAspAlaGlyLeuProGlnLysValAla 40
DB      586 GTTATCATTCAGAAATTTTCAGACATTCCTGATGCTGTTTACCAAGAAAGTTC 645
QY      41 GluLysLeuAspGlnLysValaIaGlyLeuValAlaHisSerAspLeuAspGlnArg 60
DB      646 GAAAACTGATGAAATTTTACGTTGACGGCTAGTTCACATGATGATTTAGTGAAGA 705
QY      61 AlaIleGluAlaLeuLysGluPheAsnGluAspGlyAlaLeuAlaValLeuGlnInphe 80
DB      706 GCTATTGAAGCTTTAAAGAAATTCATGAAGACGGTCATTCGACGTTCTTCAACGTTT 765
QY      81 LysAspSerAspLeuSerHisValaGlnAsnLysSerAlaPheLeuGlyValMetLys 100
DB      766 AAGACACAGATCTCTCTCATGTTTCAGAACAAAGTCCCTTTTATGTGAGATCATGAAG 825
QY      101 ThrTyArgGlnArgGlnLysGlnGlyThrLysValAlaAspSerSerLysGlyProAsp 120
DB      826 ACTTTCAGGCAAGGAGAAAAACAAGGACCAAAAGTACAGATTCATGTAAGGACAGAT 885
QY      121 GluAlaLysIleLysAlaLeuLeuGlnArgThrGlyThrThrLeuAspValThrThrGly 140
DB      886 GAGGCAAAATTTAAGCACCTTCGAAAGAACAGCTACACATTCATGATGACACATCGGA 945
QY      141 GlnArgLysThrGlyLysIleProProAspSerValTySerGlyGlnInProSerVal 160
DB      946 CAGAGGAGATGAGAGACCACTCCAGATTCGTTTATTCAGGTACGACGCTTCGTGTT 1005
QY      161 GlnTyThrGluIlePheValGlyLysIleProArgAspLeuPheGlnAspGlnLeuValPro 180
DB      1006 GGACATGAGATNTTTTGGGAAAGATCCCAAGGATCTATTGAGATGAATCTTGTTCA 1065
QY      181 LeuPheGlnLysAlaGlyProIleTPAspLeuArgLeuMetAspProLeuThrGly 200
DB      1066 TTATTGAGAAAGCTGACCTATATGAGATCTTCGCTATATGATGATTCACATCGGT 1125
QY      201 LeuAsnArgGlyTyArgLysPheValThrPheCysThrLysGluAlaIaGlnGluAlaVal 220
DB      1126 CTCATTAAGAGGTATGCGTTTGTACCTTTTGTACAAAGAACACCTCAGAGGCTGTT 1185
QY      221 LysLeuTyAsnAsnHisGlnLysLeuArgSerGlyLysHisIleGlyValCysIleSerVal 240

```

```

DB      1186 AAACCTATATATATATCATGAAATTCGTTCTGAGAAACATATGCTCTCATTCAGTT 1245
QY      241 AlaAsnAsnArgLeuPheValGlySerIleProLysSerLysThrLysGlnGlnIleLeu 260
DB      1246 GCCAACAAATAGGCTTTTGTGGGCTCTATTCCTTAAGAGTAAACCAAGAGACAGATTCCT 1305
QY      261 GlnGluPheSerLysValThrGluGlyLeuThrAspValIleLeuLysHisGlnProAsp 280
DB      1306 GAAAGATTTAGCAAGATTAACAGAGGCTTACAGAGTCATTTTATACCAACACCGAT 1365
QY      281 AspLysLysLeuAsnArgGlyPheCysPheLeuGluTyGlnAspHisLysThrAlaAla 300
DB      1366 GACAAACAAAAAACAAGAGCTTTTCCTTCTTGAATATGAAAGATCAACAAACGCTGCC 1425
QY      301 GlnAlaArgArgArgLeuMetSerGlyLysValLysValTyPGLysAsnValGlyThrVal 320
DB      1426 CAGGCAAGGCGTAGGTTAATGAGTGCTAAAGTCAGTCGAGGAGATGTTGAACTGTT 1485
QY      321 GlnTPAlAspProIleGlnAspProAspProGlnValMetAlaLysValLysValLeu 340
DB      1486 GAATGGCTGATCTATAGAGATCTGATCCTGAGGTTATGCAAAAGGTAAAGTGCCTG 1545
QY      341 PheValArgAsnLeuAlaAsnThrValThrGluGluIleLeuGlnLysSerPheSerGln 360
DB      1546 TTGTGACGACACTTCGCCAATCTGTACAGAAAGATTTTGAAAAGGCATTTAGTCAG 1605
QY      361 PheGlyLysLeuGlnArgValLysLysLeuLysAspTyArgLysPheIleHisPheAspGln 380
DB      1606 TTTGGGAAACCTGGAACGATGGAAGATTAAAGATTATGCGTTCATTCATTTTGTAGAG 1665
QY      381 ArgAspGlyAlaValLysAlaMetGlnGluMetAsnGlyLysAspLeuGlnGlyLysLeu 400
DB      1666 CAGAGTGCTGCTCAAGGCTATGAGAAAGAAATGAGAAAGACTTGAGAGGAGAAAT 1725
QY      401 IleGluIleValPheAlaLysProProAspGlnLysArgLysGlnArgLysAlaGlnArg 420
DB      1726 ATTGAATGTTTTCGCCAAGCACCAAGTCAAGAAAGAAAGAAAGAAAGCTCAGAG 1785
QY      421 GlnAlaLysAsnGlnMetTyArgAspTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 440
DB      1786 CAGGACGCAAAATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1845
QY      441 ProProThrArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArg 460
DB      1846 CCTCCAAAGAGGTGAGAGGCGGTGAGAGGTGAGAGGTGATGATGATGATGATGATGAT 1905
QY      461 TyArgLysAspTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 480
DB      1906 TATGATATGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1965
QY      481 TyArgLysAspProTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 500
DB      1966 TATGAAGATCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2025
QY      501 ArgGlyAlaArgGlyAlaAlaProSerArgGlyArgGlyAlaAlaProProArgGlyArg 520
DB      2026 AAGAGGAGCAAGGGGTGCTCTCATCCAGAGGTCTGAGGGGTGCTCTCCCGCGGTAGA 2085
QY      521 AlaGlyTySerGlnArgGlyLysIleProGlnSerAlaArgGlyValArgGlyAlaArgGly 540
DB      2086 GCCGCTTATTCACAGAGAGAGGTCTTGATGACAGAGAGGCGTTCAGAGTCTGAGAGGA 2145
QY      541 GlnAlaGlnGlnGlnArgGlyArgGlyGlyLysGlyValGlnAlaGly 556
DB      2146 GGTGCCCAACAAACAAAGAGCGCGCGGCTGACGTGTCGAGGCGGTGCG 2193
RESULTE 6
AF093821      2373 bp      mRNA      linear      ROD 04-OCT-1998
LOCUS      Mus musculus RRM RNA binding protein GRV-RBP mRNA, complete cde.
DEFINITION
ACCESSION      AF093821
VERSION      AF093821.1 GI:3694985

```











[illegible]

RESULT	9		
LOCUS	AY034482		
DEFINITION	AY034482 Homo sapiens hnRNP Q2 mRNA, complete cds.	1767 bp	linear
ACCESSION	AY034482		
VERSION	AY034482.1	GI:15809587	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		

REFERENCE	1 (bases 1 to 1767)
AUTHORS	Mourtelatos,Z., Abel,L., Yong,J., Kataoka,N. and Dreyfuss,G.
TITLE	SNM interacts with a novel family of hnRNP and spliceosomal proteins
JOURNAL	EMBO J. 20 (19), 5443-5452 (2001)
MEDLINE	21458434
PUBMED	11574476
REFERENCE	2 (bases 1 to 1767)
AUTHORS	Mourtelatos,Z.
TITLE	Direct Submission
JOURNAL	Submitted (11-MAY-2001) HEMI/ Biochemistry, University of Pennsylvania, 330 CRB, 415 Curie Blvd., Philadelphia, PA 19104, USA
FEATURES	Location/Qualifiers
source	1..1767
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
CDS	1..1767

```

      1
      2
      3
      4
      5
      6
      7
      8
      9
     10
     11
     12
     13
     14
     15
     16
     17
     18
     19
     20
     21
     22
     23
     24
     25
     26
     27
     28
     29
     30
     31
     32
     33
     34
     35
     36
     37
     38
     39
     40
     41
     42
     43
     44
     45
     46
     47
     48
     49
     50
     51
     52
     53
     54
     55
     56
     57
     58
     59
     60
     61
     62
     63
     64
     65
     66
     67
     68
     69
     70
     71
     72
     73
     74
     75
     76
     77
     78
     79
     80
     81
     82
     83
     84
     85
     86
     87
     88
     89
     90
     91
     92
     93
     94
     95
     96
     97
     98
     99
    100
    101
    102
    103
    104
    105
    106
    107
    108
    109
    110
    111
    112
    113
    114
    115
    116
    117
    118
    119
    120
    121
    122
    123
    124
    125
    126
    127
    128
    129
    130
    131
    132
    133
    134
    135
    136
    137
    138
    139
    140
    141
    142
    143
    144
    145
    146
    147
    148
    149
    150
    151
    152
    153
    154
    155
    156
    157
    158
    159
    160
    161
    162
    163
    164
    165
    166
    167
    168
    169
    170
    171
    172
    173
    174
    175
    176
    177
    178
    179
    180
    181
    182
    183
    184
    185
    186
    187
    188
    189
    190
    191
    192
    193
    194
    195
    196
    197
    198
    199
    200
    201
    202
    203
    204
    205
    206
    207
    208
    209
    210
    211
    212
    213
    214
    215
    216
    217
    218
    219
    220
    221
    222
    223
    224
    225
    226
    227
    228
    229
    230
    231
    232
    233
    234
    235
    236
    237
    238
    239
    240
    241
    242
    243
    244
    245
    246
    247
    248
    249
    250
    251
    252
    253
    254
    255
    256
    257
    258
    259
    260
    261
    262
    263
    264
    265
    266
    267
    268
    269
    270
    271
    272
    273
    274
    275
    276
    277
    278
    279
    280
    281
    282
    283
    284
    285
    286
    287
    288
    289
    290
    291
    292
    293
    294
    295
    296
    297
    298
    299
    300
    301
    302
    303
    304
    305
    306
    307
    308
    309
    310
    311
    312
    313
    314
    315
    316
    317
    318
    319
    320
    321
    322
    323
    324
    325
    326
    327
    328
    329
    330
    331
    332
    333
    334
    335
    336
    337
    338
    339
    340
    341
    342
    343
    344
    345
    346
    347
    348
    349
    350
    351
    352
    353
    354
    355
    356
    357
    358
    359
    360
    361
    362
    363
    364
    365
    366
    367
    368
    369
    370
    371
    372
    373
    374
    375
    376
    377
    378
    379
    380
    381
    382
    383
    384
    385
    386
    387
    388
    389
    390
    391
    392
    393
    394
    395
    396
    397
    398
    399
    400
    401
    402
    403
    404
    405
    406
    407
    408
    409
    410
    411
    412
    413
    414
    415
    416
    417
    418
    419
    420
    421
    422
    423
    424
    425
    426
    427
    428
    429
    430
    431
    432
    433
    434
    435
    436
    437
    438
    439
    440
    441
    442
    443
    444
    445
    446
    447
    448
    449
    450
    451
    452
    453
    454
    455
    456
    457
    458
    459
    460
    461
    462
    463
    464
    465
    466
    467
    468
    469
    470
    471
    472
    473
    474
    475
    476
    477
    478
    479
    480
    481
    482
    483
    484
    485
    486
    487
    488
    489
    490
    491
    492
    493
    494
    495
    496
    497
    498
    499
    500
    501
    502
    503
    504
    505
    506
    507
    508
    509
    510
    511
    512
    513
    514
    515
    516
    517
    518
    519
    520
    521
    522
    523
    524
    525
    526
    527
    528
    529
    530
    531
    532
    533
    534
    535
    536
    537
    538
    539
    540
    541
    542
    543
    544
    545
    546
    547
    548
    549
    550
    551
    552
    553
    554
    555
    556
    557
    558
    559
    560
    561
    562
    563
    564
    565
    566
    567
    568
    569
    570
    571
    572
    573
    574
    575
    576
    577
    578
    579
    580
    581
    582
    583
    584
    585
    586
    587
    588
    589
    590
    591
    592
    593
    594
    595
    596
    597
    598
    599
    600
    601
    602
    603
    604
    605
    606
    607
    608
    609
    610
    611
    612
    613
    614
    615
    616
    617
    618
    619
    620
    621
    622
    623
    624
    625
    626
    627
    628
    629
    630
    631
    632
    633
    634
    635
    636
    637
    638
    639
    640
    641
    642
    643
    644
    645
    646
    647
    648
    649
    650
    651
    652
    653
    654
    655
    656
    657
    658
    659
    660
    661
    662
    663
    664
    665
    666
    667
    668
    669
    670
    671
    672
    673
    674
    675
    676
    677
    678
    679
    680
    681
    682
    683
    684
    685
    686
    687
    688
    689
    690
    691
    692
    693
    694
    695
    696
    697
    698
    699
    700
   
```

<b>Alignment Scores:</b>	
Pred. No.:	1,45e-118
Score:	2705.50
Percent Similarity:	92.81%
Best Local Similarity:	92.45%
Query Match:	90.73%
DB:	
	9
	Gaps:
	1

Qy	1	MeLiaThrgIuHlSvaIsnGIYaSnGIYThrGIuGIuProMeAspThrThSerAla	20
Db	1	ATGGCTACAGAAACATGTTAAATGGAAATGTAICTAGAGGCCCATGGATACACTTCTTCCA	60
Qy	21	VallIeHsSerGIuAsnPhcGIInThrLeuLeuAspAlaGIYLeuProGInLYsValAla	40
Db	61	GTTATCCATTACAGAAATTTTTCACACATTCGTTATGTCGTTTATACACAGAAAGTTGCT	120
Qy	41	GIuLYsLeuAspGIuIeTYrValAlaGILYeuValAlaHIsSerAspLeuAspGIuArg	60
Db	121	GAAAAACCTAGTGAATAATTTCAGTCGAGGCTTAAGTGACATAGTGAATTITAGATGAAGA	180
Qy	61	AlaIleGuaIaLeuLYsGIuPhcAsnGIuAspGIYAlaLeuAlaValLeuGInGIuPhe	80
Db	181	GCATTTCGAAGCTTTAAAAAGAAATTCATAGAGACGGTCGATTCGACAGTTCCTCAACAGTTT	240
Qy	81	LYsAspSerAspLeuSerHisValGIuAsnLYsSerLAlaPheLeuCYsGIYValMetLYs	100
Db	241	AAAGACAGTGAATCTCTCTCATGTTCAAGACAAAGATGCTTTTATGTGGAGTATAGAAG	300
Qy	101	ThrTYrArgGIuArgGIuLYsGIuLYsThrLYsValAlaAspSerSerLYsGIYProAsp	120
Db	301	ACTTACAGGCGAGAGAAAAACAAGGCCAAAGTACAGATTCTTAGTAAAGACCAAGAT	360
Qy	121	GIuAlaLYsIleLYsAlaLeuLeuGIuArgThrGILYrThrLeuAspValThrThGILY	140
Db	361	GAGCGAAAAATTTAAGGCACTCTTGAAAAAGAACAGGCTTACACACTTGTGATGTGACCACTGGA	420

QY 141 GlnArgLysTyrGlyGlyProProAspSerValTyrSerGlyGlnGlnProSerVal 160  
 DB 421 CAGAGAGAGTATGAGAGACCACTCCAGATTCCGTTTACAGGTCAGACGCTTCTTT 480  
 QY 161 GlyThrGluIlePheValGlyLysIleProArgAspLeuPheGluAspGluLeuValPro 180  
 DB 481 GGCCTGATGATATTTGTGGGAAAGATCCCAAGAGATCTATTTGAGAGTGAAGCTTGTTCA 540  
 QY 181 LeuPheGluLysValIleGlyProIleThrAspLeuArgLeuMetMetAspProLeuThrGly 200  
 DB 541 TTATTTGAGAAAGCTGAGCCTATATGAGATCTTCTGTTAAGATGAGATCCACTACGTGGT 600  
 QY 201 LeuAsnArgGlyTyrAlaPheValThrPheCysThrLysGlnAlaAlaGlnGluAlaVal 220  
 DB 601 CTCAATAGAGTTATGCTTGTCTTCTTTGTACAAAGAGAGCTCAGAGGCTGTT 660  
 QY 221 LysLeuTyrAsnAsnHisGluIleArgSerGlyLysHisIleGlyValCysIleSerVal 240  
 DB 661 AACTGATATATATCATGAATTCGTTCTGAAAAACATATGGGTCTGCATCTCAGTT 720  
 QY 241 AlaAsnAsnArgLeuPheValGlySerIleProLysSerLysThrLysGlnGlnIleLeu 260  
 DB 721 GCCAACAATAGCTTTTGTGGCTTATTTCTTAAGATGAAACCAAGAACAGATTTCTT 780  
 QY 261 GluGluPheSerLysValThrGluGlyLeuThrAspValIleLeuTyrHisGlnProAsp 280  
 DB 781 GAAGATTTTACCAAGTACAGAGGCTTTACAGCGCATTTTATACCAACAGATTTCTT 840  
 QY 281 AspLysLysLysAsnArgGlyPheCysPheLeuGluTyrGluAspHisLysThrAlaAla 300  
 DB 841 GACAGAGAAAAACAGAGAGCTTTGCTTTCTTGAAATGAAATCACAACAGAGCTGCC 900  
 QY 301 GlnAlaArgArgLeuMetSerGlyLysValLysValTyrGlyValAsnValGlyThrVal 320  
 DB 901 CAG----- 903  
 QY 321 GluTyrAlaAspProIleGluAspProAspProGluValMetAlaLysValLysValLeu 340  
 DB 904 -----GTAAAAGTCTG 915  
 QY 341 PheValArgAsnLeuAlaAsnThrValThrGluGluIleLeuGluLysSerPheSerGln 360  
 DB 916 TTGTAGGCAACTTGGCAATACGTAAACAGAGATTTTGAAGAGCGCATTTTATGTCAG 975  
 QY 361 PheGlyLysLeuGluArgValLysLeuLysLysAspTyrAlaPheIleHisPheAspGlu 380  
 DB 976 TTTCGAAACTGGAACGAGTAAAGATTAAGATTAAGTTCATTCATTTGATGAG 1035  
 QY 381 ArgAspGlyAlaValLysAlaMetGluGluMetAsnGlyLysAspLeuGluGluAsn 400  
 DB 1036 CGAGATGCTGCTCAAGGCTATGGAAGATGAATGSCAAAGACTTGAGGAGGAAAT 1095  
 QY 401 IleGluIleValPheAlaLysProProAspGlnLysArgLysGluArgLysAlaArg 420  
 DB 1096 ATTCAAATTTGTTTGGCAAGCCACAGATCAGAAAGAAAGAAAGAAAGCTCAGAGG 1155  
 QY 421 GlnAlaAlaLysAsnGlnMetTyrAspAspTyrTyrTyrTyrProPheHisMetPro 440  
 DB 1156 CAGACGCAAAAAATCAATATGACGATTAATCACTATTAATGTCACCTCAATATGCC 1215  
 QY 441 ProProThrArgGlyArgGlyArgGlyGlyValArgGlyGlyTyrProProAspTyr 460  
 DB 1216 CCTCCAAACAAGAGTCCAGGCGTGGAGGTAGAGTGTATGATATCCCTCCAGATTTAT 1275  
 QY 461 TyrGlyTyrGluAspTyrTyrAspTyrTyrGlyTyrAspTyrHisAsnTyrArgGlyGly 480  
 DB 1276 TATGATATGAAAGATTAATGATTAATGATTAATGATTAATCACTAATCGTGGAGA 1335  
 QY 481 TyrGluAspProTyrTyrGlyTyrGluAspPheGlnValGlyAlaArgGlyArgGlyGly 500  
 DB 1336 TATTAAGATTCATCTATGTTATGAAAGATTTTCAAGTTGAGGCTGAAGAAAGGGGTGGT 1395  
 QY 501 ArgGlyAlaArgGlyAlaAlaProSerArgGlyArgGlyAlaAlaProProArgGlyArg 520

DB 1396 AGAGAGCAAGGGGTGCTCTCCATCAGAGTCTGGGGCTCTCTCCCGCGTAA 1455  
 QY 521 AlaglyTyrSerGlnArgGlyGlyProGlySerAlaArgGlyValaArgGlyAlaArgGly 540  
 DB 1456 GCCGTTATTCACAGAGAGAGGCTCTGATCAGCAAGAGCGCTTCAGGTGGAGGA 1515  
 QY 541 GlyAlaGlnGlnGlnArgGlyArgGlyGlyValGluAlaGly 556  
 DB 1516 GGTGCCCAACAAAGAGCGCGGAGTACGTGTCCGAGGGGTGGC 1563  
 RESULT 10  
 HSJ775C13  
 LOCUS  
 DEFINITION  
 Human DNA sequence from clone RP4-775C13 on chromosome 20p12.1-13. Contains an hnRNP R (heterogenous nuclear riboprotein R, RRM rna binding protein GRP-RBP) pseudogene, ESTs, STSs and GSS, complete sequence.  
 ACCESSION  
 AL109618.1 GI:5668665  
 VERSION  
 HTG; GRP-RBP; hnRNP; RRM.  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Direct Submision  
 Submitted (02-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 requester: clonerequest@sanger.ac.uk  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., paired quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep  
 This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr20  
 RP4-775C13 is from the library RPCR-4 constructed by the group of Plier de Jong. For further details see  
 http://www.chori.org/bacpac/home.htm  
 VECTOR: pCYPAC2  
 This sequence is the entire insert of clone RP4-775C13.  
 FEATURES  
 source  
 1..103819  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="p12.1-13"  
 /clone="RP4-775C13"  
 /clone\_1fb="RPCR-4"  
 repeat\_region  
 1..236  
 /note="AluY repeat: matches 1..236 of consensus"  
 repeat\_region  
 293..562  
 /note="L1Mc repeat: matches 1844..1624 of consensus"  
 repeat\_region  
 819..1471  
 /note="L1M3c repeat: matches 969..1634 of consensus"  
 repeat\_region  
 2177..2588

```

repeat_region /note="LIM3 repeat: matches 14. .431 of consensus"
                2593. .2652
                /note="MUTIF repeat: matches 474. .533 of consensus"
                2701. .3176
                /note="238 copies 2 mer tt 54% conserved"
                2814. .2978
                /note="15 copies 11 mer 66% conserved"
                2996. .3196
                /note="67 copies 3 mer tcc 56% conserved"
                3031. .3184
                /note="14 copies 11 mer 63% conserved"
                3296. .4082
                /note="LIM4 repeat: matches 3350. .4128 of consensus"
                4197. .4315
                /note="LIM4 repeat: matches 3021. .3136 of consensus"
                4316. .4738
                /note="MSTA repeat: matches 1. .426 of consensus"
                4339. .5118
                /note="LIM4 repeat: matches 2626. .3021 of consensus"
                5146. .5288
                /note="match: GSS: Em:B90786"
                5345. .6619
                /note="LIM4 repeat: matches 3863. .5135 of consensus"
                6882. .7139
                /note="129 copies 2 mer gg 58% conserved"
                7351. .8891
                /gene="du775C13.1"
                /note="du775C13.1" (hnRNP R (heterogenous nuclear
                riboprotein R, RRM rna binding protein GRY-RBP)
                pseudogene)
                match: cDNAs: Em:AF000364 Em:AF037448 Em:AF093821
                match: ESTs: Em:AA411147 Em:W38408 Em:AT65128 Em:AV147169
                Em:AA481148 Em:AA307374 Em:TB89820 Em:AA677426 Em:TB9909
                Em:AT688219 Em:AA1472757 Em:AT548565 Em:AT134944
                Em:AA912195 Em:AA094744 Em:AT084864 Em:H34275 Em:AA403140
                Em:AT289939 Em:D76738 Em:AA146046 Em:TB9242 Em:AA075580
                Em:AA561931 Em:AA647246 Em:AV115211 Em:AA356549
                Em:AT581194 Em:H32663 Em:AT566089 Em:AA207215 Em:AA216046
                Em:TB5450 Em:AA370336 Em:AA65413 Em:W39501 Em:MG1969
                Em:AA164877 Em:AA403169 Em:AA65344 Em:NB6532 Em:R29647
                Em:AA298076 Em:AT625087 Em:AA164209 Em:AA164208
                Em:AA563205 Em:AA066208 Em:AA045159 Em:AA524815
                Em:AA544174 Em:AA188195 Em:AA645286 Em:AA164876
                Em:AT267372 Em:AA130302 Em:AA645286 Em:AA662966
                Em:AA130625 Em:AA232633 Em:AA242157 Em:AA130793
                Em:AA654252 Em:AA306863 Em:AA186327 Em:AA153225
                Em:AA418952 Em:TB9722 Em:AA629942 Em:AA95153 Em:AA594467
                Em:AV14113 Em:AT469082 Em:AT263768 Em:AA124310
                Em:AT072424 Em:AA313602
                match: proteins: Tr:O23093 Tr:O24188 Tr:O64985 Tr:O01671
                Sw:P22626 Sw:P38159 Tr:O88569 Tr:O60506 Tr:O88991
                Tr:O22385 Tr:O76494 Tr:O50602 Tr:O71309 Tr:O61474
                Tr:O22390 Wp:CE021230 Tr:O97560 Sw:P48810 Tr:O43347
                Tr:O24668 Wp:CE03230 Tr:O76181 Tr:P48158 Tr:O43390
                Tr:O17310 Sw:O39020 Sw:P51968 Tr:O91579 Wp:CE17059
                Sw:P49314 Sw:O13247 Tr:O04240 Tr:O15717 Tr:O61954
                Sw:O08473 Wp:P39697 Tr:O23288 Tr:O43349 Tr:O81987
                Sw:P51991 Wp:CE00983 Tr:O08937 Tr:P90699 Tr:O13550
                Sw:Q01560 Tr:Q24409 Tr:P70372 Tr:O61374 Tr:Q39953
                Tr:P70807 Wp:CE02197 Tr:O14979"
                /pseudocodon_start=1
                /evidence=not_experimental
                7549. .8004
                /gene="du775C13.1"
                /note="match: GSS: Em:AA0434110"
                7614. .7732
                /gene="du775C13.1" ✓
                /note="match: GSS: Em:AA0536515"
                complement(6069. .8562)
                /note="match: STS: Em:TB97722"

```

```

misc_feature      complement (8476. .9063)
                  /note="match: GSS: Em:AQ008463"
repeat_region     8721. .8816
                  /note="32 copies 3 mer att 65% conserved"
repeat_region     9341. .10347
                  /note="L1PA16 repeat: matches 5128. .6150 of consensus"
repeat_region     10381. .10517
                  /note="L1M4 repeat: matches 2211. .2348 of consensus"
repeat_region     11105. .11201
                  /note="HAI1 repeat: matches 675. .775 of consensus"
repeat_region     11261. .11455
                  /note="L2 repeat: matches 645. .869 of consensus"
repeat_region     11456. .11765
                  /note="A1uSx repeat: matches 1. .307 of consensus"
repeat_region     11766. .11863
                  /note="L2 repeat: matches 555. .645 of consensus"
repeat_region     12356. .12519
                  /note="L2 repeat: matches 298. .464 of consensus"
repeat_region     12520. .12615
                  /note="L1HS repeat: matches 6051. .6146 of consensus"
repeat_region     13257. .13396
                  /note="70 copies 2 mer ga 69% conserved"
repeat_region     13442. .13893
                  /note="MER54B repeat: matches 1. .498 of consensus"
repeat_region     13894. .13931
                  /note="19 copies 2 mer ac 89% conserved"
repeat_region     14475. .14910
                  /note="L1M4C repeat: matches 5707. .6148 of consensus"
repeat_region     14560. .15480
                  /note="L1M4 repeat: matches 4030. .4591 of consensus"
repeat_region     15498. .16282
                  /note="L1PA4 repeat: matches 5361. .6146 of consensus"
repeat_region     16286. .16427
                  /note="L1PA4 repeat: matches 5213. .5358 of consensus"
misc_feature      16428. .16958
                  /note="match: GSS: Em:AQ427784"
repeat_region     17470. .18240
                  /note="L1PB1 repeat: matches 5277. .6151 of consensus"
repeat_region     18383. .18706
                  /note="A1uY repeat: matches 1. .309 of consensus"
repeat_region     18793. .18837
                  /note="15 copies 3 mer att 75% conserved"
repeat_region     19213. .19509
                  /note="A1uB repeat: matches 3. .298 of consensus"
misc_feature      complement (join(19771. .19905,51843. .51912))
                  /note="match: GSS: Em:AQ3778847"
repeat_region     22426. .22613
                  /note="MER53 repeat: matches 4. .188 of consensus"
repeat_region     23385. .23561
                  /note="MER5B repeat: matches 1. .178 of consensus"
repeat_region     23821. .23871
                  /note="L2 repeat: matches 2445. .2496 of consensus"
repeat_region     23899. .24082
                  /note="MER5A repeat: matches 1. .181 of consensus"
repeat_region     24363. .24676
                  /note="A1uSx repeat: matches 3. .312 of consensus"
repeat_region     25302. .25481
                  /note="L1R1G repeat: matches 9. .187 of consensus"
repeat_region     25486. .25537
                  /note="26 copies 2 mer ct 86% conserved"
repeat_region     25564. .25591
                  /note="L1R1G repeat: matches 243. .270 of consensus"
repeat_region     26477. .26762
                  /note="143 copies 2 mer tc 56% conserved"
repeat_region     26510. .26718
                  /note="19 copies 11 mer 66% conserved"
repeat_region     26836. .27041

Alignment Scores:
Pred. No.:      2.26e-115      Length:      103819
Score:          2676.00        Matches:      537
Percent Similarity: 95.93%      Consensivity: 537
Best Local Similarity: 95.04%   Mismatches:  18

```



```

Center project name: BA190P17
----- Summary Statistics -----
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 203924 bases at least Q40
Consensus quality: 207400 bases at least Q30
Consensus quality: 203306 bases at least Q20
Insert size: 21159; sum-of-contigs
Quality coverage: 7.54x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 48415: contig of 48415 bp in length
* 48416 48515: gap of 100 bp
* 48516 82643: contig of 34128 bp in length
* 82644 82743: gap of 100 bp
* 82744 89101: contig of 6358 bp in length
* 89102 89201: gap of 100 bp
* 89202 127367: contig of 38166 bp in length
* 127368 127467: gap of 100 bp
* 127468 169067: contig of 41600 bp in length
* 169068 169167: gap of 100 bp
* 169168 171259: contig of 2092 bp in length
* 171260 171359: gap of 100 bp
* 171360 173612: contig of 2253 bp in length
* 173613 173712: gap of 100 bp
* 173713 187445: contig of 13733 bp in length
* 187446 187545: gap of 100 bp
* 187546 191084: contig of 3539 bp in length
* 191085 191184: gap of 100 bp
* 191185 194709: contig of 3525 bp in length
* 194710 194809: gap of 100 bp
* 194810 197824: contig of 3015 bp in length
* 197825 197924: gap of 100 bp
* 197925 200623: contig of 2699 bp in length
* 200624 200723: gap of 100 bp
* 200724 203433: contig of 2710 bp in length
* 203434 203533: gap of 100 bp
* 203534 206523: contig of 2990 bp in length
* 206524 206623: gap of 100 bp
* 206624 208875: contig of 2252 bp in length
* 208876 208975: gap of 100 bp
* 208976 212659: contig of 3684 bp in length.
*
Location/Qualifiers
1. 212659
/organism="Homo sapiens"
/db xref="taxon:9606"
/chromosome="13"
/clone="RP11-190P17"
/clone_1fb="RPCT-11.1"
1. 48415
/note="assembly_fragment:03097
fragment_chain:1"
misc_feature
48516..82643
/note="assembly_fragment:01394
fragment_chain:1"
misc_feature
82744..89101
/note="assembly_fragment:01816
fragment_chain:1"
misc_feature
89202..127367
/note="assembly_fragment:03032
fragment_chain:2"
misc_feature
127468..169067
/note="assembly_fragment:00218
fragment_chain:2"
misc_feature
169168..171259

```

```

misc_feature /note="assembly_fragment:00015"
171360..173612
/note="assembly_fragment:00477"
misc_feature 173713..187445
/note="assembly_fragment:00780"
misc_feature 187546..191084
/note="assembly_fragment:01534"
misc_feature 191185..194709
/note="assembly_fragment:01673"
misc_feature 194810..197824
/note="assembly_fragment:02018"
misc_feature 197925..200623
/note="assembly_fragment:02519"
misc_feature 200724..203433
/note="assembly_fragment:02839"
misc_feature 203534..206523
/note="assembly_fragment:02989"
misc_feature 206624..208875
/note="assembly_fragment:03934"
misc_feature 208976..212659
/note="assembly_fragment:04063"
BASE COUNT 61246 a 41295 c 42666 g 65947 t 1505 others
ORIGIN
Alignment Scores:
Pred. No.: 4,71e-115 Length: 212659
Score: 2676.00 Matches: 537
Percent Similarity: 95.93% Conservatve: 5
Best Local Similarity: 95.04% Mismatches: 18
Query Match: 89.74% Indels: 9
DB: 2 Gaps: 2
US-09-821-687-4 (1-561) x AL390737 (1-212659)
QY 1 MetAlaThrGluHisValAsnGlyAsnGlyThrGluGluProMetAspThrSerAla 20
DB 83107 ATGGCTACGAGAACATTTATGAAATGATCTGAAGGCCATGATCTCTTCTGCA 83166
QY 21 ValIleHisSerGluAsnDhncGlnThrLeuLeuAspAlaGlyLeuProGlnValAla 40
DB 83167 GTTATCATTCAGAAATTTTCAGACATTCGTTGATGCTGTTTATCAGAAAGTGT 83226
QY 41 GluValLeuAsnGluIleLeuValAlaGlyLeuValAlaHisSerAspLeuAspGluAla 60
DB 83227 GAAAACTGAGTGAATTTACCTGCGAGGCTGATGCAATGATGCAATGATGAAAGA 83286
QY 61 AlaIleGluAlaLeuValAsnGluProAsnGluValAlaLeuAlaValLeuGlnGlnPhe 80
DB 83287 GCTATTGAGCTTTAAAGAAATTCATGAAAGAGGATGATGCGATGCTTCAACAGTT- 83345
QY 81 LysAspSerAspLeuSerHisValGlnAsnLysSerAlaPheLeuGlyValMetLys 100
DB 83346 AAGAGACAGATCTCTCTGCTGTTAGAAACAAAGAGCTTTTATGTCGATCATGAA 83405
QY 101 ThrTyrArgGlnArgGluLysGlnGlyThrLysValAlaAspSerSerLysGlyProAsp 120
DB 83406 CCTTATAGGCAAGAGAAACAAAGGACCAAGATGACGATTTAGTAAAGACAGAT 83465
QY 121 GluAlaLysIleLeuValAlaLeuLeuGluArgThrGlyTyrThrLeuAspValThrThrGly 140
DB 83466 GAGGCAAAATTTAAGGCACTCTTGAAGAGAGGCTGACACTGATGATGACCACTGGA 83525
QY 141 GlnArgLysTyrGlyGlyProProProAspSerValTyrSerGlyGlnGlnProSerVal 160
DB 83526 CAGAGGAATATGAGAGACCACTCCAGATTCGTTTATTCAGGTCAGAGGCTCTGTT 83585
QY 161 GlyThrGluIlePheValGlyLysIleProArgAspLeuPheGluAspGluLeuValPro 180
DB 83586 GGCACCTGAGATATTTGTGGAAAGATCCCAAGAGATCTATTTGAGAGATCACTTGTCCA 83645
QY 181 LeuPheGluLysAlaGlyProIleTyrAspLeuAspGluMetMetAspProLeuThrGly 200
DB 83646 TTATTGGAAGAGCTGCTCTTATGAGATCTTCCTTAATGATGATCCACTGCTGT 83705

```



QY 201 LeuAsnArgGlyTyrAlaPheValThrPheCysThrLeuGluAlaAlaGlnGluAlaVal 220  
 DB 83706 CTCATAGAGGTATAGCTGTGTGCACTTTGTGACAAA--GAGGCTGGGAGGCGTGT 83762  
 QY 221 -LysLeuTyrAsnAsnHisGlnIleArgSerGlyLysHisIleGlyValCysIleSerVa 240  
 DB 83763 AAAACGTATATATATCATGAAATTCGTTCTGAAAACACATTCGTGTCTGCATCTCAGT 83822  
 QY 240 LAlaAsnAsnArgLeuPheValGlySerIleProLysSer-LysThrLysGluGlnIleL 260  
 DB 83823 TGCACAAATAGAGCTTTTGTGGGCTCTATCTTAAGAGTAAACCAAGACAAACAGATTC 83882  
 QY 260 euGluGluPheSer-LysValThrGluGlyLeuThrAspValIleLeuTyrHisGlnPro 279  
 DB 83883 TTGAAGAAATTTAGCAAAAGTACAGAGGCTTTCACAGCTCATTTTATACCAACCG 83942  
 QY 280 AspAspLysLysLysAsnArgGlyPheCysPheLeuGluTyrGluAspHisLysThrAla 299  
 DB 83943 GATGACAAAGAAAACAA--GGCTTTTGTCTTCTGAAATGAAAGTCAACAAACAGCT 84000  
 QY 300 AlAGlnAlaArgArgGluGluMetSerGlyLysValLysValITPGLysAsnValGlyThr 319  
 DB 84001 GCCCAGGCAAGGCGTAGGTATAGTGAAGTCAAGTCAAGCTCTGGGGGAATGTTGAACT 84060  
 QY 320 ValGluTrrAlaAspProIleGluAspProAspProGluValMetAlaLysValLysVal 339  
 DB 84061 GTTAAAGGCTGATCCATGAGAGATCTGATCTGAGTTAGGCAAGGTAAGAGTGA 84120  
 QY 340 LeuPheValArgAsnLeuAlaAsnThrValThrGluGluIleLeuGluLysSerPheSer 359  
 DB 84121 CTGTTTGAAGCAACTTGGCCAAATACGTACGGAAGAGATTTTGAAGAAAGCATTTAGT 84180  
 QY 360 GlnPheGlyLysLeuGluLysValLysLysLeuLysAspTyrAlaPheIleHisPheAsp 379  
 DB 84181 CAGTTTGGGAACTGGAATGAGTGAAGAGTAAAGATTTTCATTCATTCATTTTGTGT 84240  
 QY 380 GluArgAspGlyAlaValLysValMetGluGluMetAsnGlyLysAspLeuGluGlu 399  
 DB 84241 GAGGAGATGAGTGTCTGACAGGCTATGAGAGATGATGATGATGATGATGATGATGAT 84300  
 QY 400 AsnIleGluIleValPheAlaLysPheProAspGlnLysArgLysGluArgLysAlaGln 419  
 DB 84301 AATTTGAAATTTGTTTGGCCAAACCAATACGAAAGAAAGAA--AAAGCTCAG 84358  
 QY 420 ArgGlnAlaAlaLysAsnGlnMetTyrAspAspTyrTyrTyrTyrTyrProHisMet 439  
 DB 84359 AGGCACAGCAGCAAAATCAATGATGACGATTACTATTATGCTCCACTTATATG 84418  
 QY 440 ProProProThrArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGly 459  
 DB 84419 CCCCCCTCAACAGAGATGAGAGGCGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 84478  
 QY 460 TyrTyrGlyTyrGluAspTyrTyrAspTyrTyrGlyTyrAspTyrHisAsnTyrArgGly 479  
 DB 84479 TATTATGATATGAGAGTATATATGATATATATGATATATGATATATGATATATGAT 84538  
 QY 480 GlyTyrGluAspProTyrTyrGlyTyrGluAspPheGlnValGlyAlaArgGlyArgGly 499  
 DB 84539 GGAATATGAGATCACTATGATGATGATGATGATGATGATGATGATGATGATGATG 84597  
 QY 500 GlyArgGlyAlaArgGlyAlaAlaProSerArgGlyArgGlyAlaAlaProProArgGly 519  
 DB 84598 GGTAG 84657  
 QY 520 ArgAlaGlyTyrSerGlnArgGlyGlyProGlySerAlaArgGlyValArgGlyAlaArg 539  
 DB 84658 AGAGCCGCTTATTCACAG 84717  
 QY 540 GlyGlyAlaGlnGlnGlnArgGlyArgGly--GlyLysGlyValGluAlaGlyProAsp 558  
 DB 84718 GGAAGTGTCCCAACAAAG 84777

QY 559 LeuLeuGln 561  
 DB 84778 CTCATTACAA 84786  
 RESULT 12  
 AF408434  
 LOCUS  
 DEFINITION Mus musculus RRM RNA binding protein NSAP1 (Nsap1) mRNA, partial  
 cds AF408434  
 AF408434.1 GI:15986734  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS Zhou, M. and Raschke, W.C.  
 TITLE Mus musculus RRM RNA binding protein, NSAP1  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 2232)  
 AUTHORS Zhou, M. and Raschke, W.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-AUG-2001) Immunology, Sidney Kimmel Cancer Center,  
 10835 Altman Row, San Diego, CA 92121, USA  
 FEATURES  
 source  
 location/Qualifiers  
 1..2232  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 gene  
 1..2232  
 /gene="Nsap1"  
 /gene="Nsap1"  
 /db\_xref="GI:15986735"  
 /db\_xref="GI:15986735"  
 /translation="KRFNEDGALAVLQPKQSDLSHONKSAFLGGMKTYRORRQK  
 TRVADSSKQPDPAKIKALLERGTLLDTTGGQRKGGPPDSVYSGQPSVTEIFVG  
 KIPRLFDLPLLEKAGPIWDLRLMDPLGLNRGYAFVTEKFAAQEAVKLYNN  
 HEIRSGKAGICIVANNRLPVGSIKPKETKQILIEBSKVTBGLTDVILYHQPDCK  
 KNRGCFLEVEHDKIAQARRLMGSKYVWNVGTVMADPIEDPDEVMKAVLFL  
 VNLANLTBEELEKSFQFGLERVKLKDYAFIHPERDGAVKAMEBNKDLGE  
 NIEIYAPKDPDQREKRRKROQAQKQWVDVYVYGPMPPTGRRGRRGGVGYG  
 PDYTGEDYDTYTGIDYHNTKGTEDPTTGIEDFYGKRGKRGKAGAPRRGGAA  
 PRRGAGYSQRGSPSARGVARGAGAAQQRGVRGARGAGGGGAVGGRKADGYNP  
 DTKRQRTNNQNSQPIAQOPLQGDGHSNKGSEBNQFYDTRGQOMK"  
 ORIGIN  
 BASE COUNT 689 a 387 c 551 g 605 t  
 Alignment Scores:  
 Pred. No.: 1.03e-113 Length: 2232  
 Score: 2604.00 Matches: 486  
 Percent Similarity: 99.19% Conservative: 1  
 Best Local Similarity: 98.98 Mismatches: 4  
 Query Match: 87.32% Indels: 0  
 DB: 10 Gaps: 0

US-09-821-687-4 (1-561) x AF408434 (1-2232)  
 QY 66 LysGluPheAsnGluAspGlyAlaLeuAlaValLeuGlnGlnPheLysAspSerAspLeu 85  
 DB 3 AAAGAGTTCAATAG 62  
 QY 86 SerHisValGlnAsnLysSerAlaPheLeuCysGlyValMetLysThrTyrArgGlnArg 105  
 DB 63 TCCATGTTCAAGAAACAAAGTGCCTTTTATATGAGTATGAGTAAAGCTTAAGGAGAGA 122  
 QY 106 GluLysGlnGlyThrLysValAlaAspSerSerLysGlyProAspGluAlaLysIleLys 125  
 DB 123 GAAACACAGGGGACCAAG 182  
 QY 126 AlaLeuLeuGluLysThrGlyTyrThrLeuAspValThrThrGlyGlnArgLysTyrGly 145

Db 183 GCACCTTTGGAAAGAGAGGCTACACCTGATGTGACTACAGGTACAGAGGAAGTATGGA 242  
 Qy 146 G1ProProPaaPaaSerValTyrSerG1yGlnGlnProSerValG1yThnGlnIlePhe 165  
 Db 243 GAAACCACTCCAGATCCGTTATTCAGGTACAGAGGCTTCTGTGGCAGCATGAAATTTT 302  
 Qy 166 ValG1yLysIleProAArgPaaPaaPheG1uApsG1uApsValProLeuPheG1uLysAla 185  
 Db 303 GTGGGGAAGATCCCAAGATCTCTTTGAGATAGAGCTGTGCTCATTTATTTGAGAAAGCT 362  
 Qy 186 G1yProIleTrrPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 205  
 Db 363 GACCTTATAGGATCTCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 422  
 Qy 206 AlaPheValThrPheCysThrIleValAlaAlaGlnGlnIleValAlaValIleValLeuTyrAsnAsn 225  
 Db 423 GCGTTGTCTACTTTTGTACAAAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 482  
 Qy 226 H1eG1uIleArgSerG1yLysH1eIleG1yValCysIleSerValAlaAsnAsnArgLeu 245  
 Db 483 CATGAATTCGTTCCGGGAGCAGCATTTGGTGTCTGATCTCAGTTGCCAACAATAGCCTT 542  
 Qy 246 PheValG1ySerIlePaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 265  
 Db 543 TTGTGGGCTCGATTCCTAAGAGTAAACCAAGAGCAGATTCCTTGAGGATTTAGCAAA 602  
 Qy 266 ValThnGlnIleLeuThrPaaPaaValIleLeuTyrH1eGlnProPaaPaaPaaPaaPaaPaaPaa 285  
 Db 603 GTACAGAGGCTCTCAGAGATGATCATTATTTATCCACCAACCTGATGACAGAAAGAAAC 662  
 Qy 286 ArgG1yPheCysPheLeuG1uTyrG1uApsH1ePheThraIaIaGlnIleAArgArgArg 305  
 Db 663 AGAGGCTTTGCTTCTTGATATAGAGATCACAAAACAGCTGCCAGCAGAGCTAGG 722  
 Qy 306 LeuMetSerG1yLysValIleValTyrG1yAsnValG1yThraValG1uTrrPaaPaaPaaPaa 325  
 Db 723 CTATAGTGTGTAAGTCAAGTCTGGGAAATGTTGAACTGTGAGTGGGCTGATCCT 782  
 Qy 326 H1eG1uApsPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 345  
 Db 783 ATTAAGATCTCTATCTCTGAGGATGAGCAAGGTAAGTGTGCTGTTGATCCCAACCTT 842  
 Qy 346 AlaAsnThrValThnGlnIleLeuG1uLysSerPheSerGlnPheG1yLysLeuGln 365  
 Db 843 GCCAACAAGGCTTACAGAGAAATTTTGAAGAGCTTTAGTGTGAGTGGGAACTGGAA 902  
 Qy 366 ArgValIleLysLeuLysPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 385  
 Db 903 CGAGTAAAGAGCTAAAGATTAATGCTTCATTCATTTGATGAGAGATGGTGTCTC 962  
 Qy 386 LysAlaMetG1uGlnIleMetAsnG1yLysPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 405  
 Db 963 AAGGCTATGAGAAAGAAATGTAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1022  
 Qy 406 AlaLysProPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 425  
 Db 1023 GCTAAGCAGCAATCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1082  
 Qy 426 GlnMetTyrAsnAsnTyrTyrTyrTyrG1yProProH1eMetProProProProH1eMetProPro 445  
 Db 1083 CAATGATATGATATTAATCTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1142  
 Qy 446 ArgG1yLysG1yLysG1yLysG1yLysG1yLysG1yLysG1yLysG1yLysG1yLysG1yLys 465  
 Db 1143 CGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1202  
 Qy 466 TyrTyrAsnTyrTyrG1yLysPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 485  
 Db 1203 TATTAAGATTAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1262  
 Qy 486 TyrG1yTyrG1uApsPheGlnValG1yAlaArgG1yArgG1yLysG1yLysG1yLysG1yLysG1y 505

Db 1263 TATGTTATGAGATTTTCAAGTTTGAGCTAGAGAGGGGTGTAGAGAGCAAGGGGT 1322  
 Qy 506 AlaAlaProSerArgG1yArgG1yAlaAlaProProArgG1yArgAlaG1yTyrSerGln 525  
 Db 1323 GCTGTCCATCCAGAGGTCTGTGGGCTGTCTCTCCCGGTGTAGAGCCGGTATTTACAG 1382  
 Qy 526 ArgG1yG1yProG1ySerAlaArgG1yValArgG1yAlaArgG1yG1yLysGlnGlnGln 545  
 Db 1383 AGAGAG 1442  
 Qy 546 ArgG1yArgG1yG1yLysG1yValG1uAlaG1y 556  
 Db 1443 AGAG 1475  
 RESULT 13  
 AX011753  
 LOCUS AX011753 3101 bp DNA linear PAT 06-SEP-2000  
 DEFINITION Sequence 156 from Patent WO955858.  
 ACCESSION AX011753  
 VERSION AX011753.1 GI:9998277  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 3101)  
 AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and  
 Pilsbry,C.  
 TITLE Human nucleic acid sequences obtained from pancreas tumor tissue  
 JOURNAL Patent: WO 955858-A 156 04-NOV-1999;  
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN  
 BERND (DE); ROSENTHAL ANDRE (DE); METZEN GBS FUER GENOMFORSCHUN  
 (DE); PILSKY CHRISTIAN (DE)  
 FEATURES  
 source 1..3101  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 980 a 540 c 670 g 911 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 8,486-109 Length: 3101  
 Score: 2502.00 Matches: 476  
 Percent Similarity: 92.98% Conservative: 1  
 Best Local Similarity: 92.79% Mismatches: 0  
 Query Match: 83.90% Indels: 36  
 DB: Gaps: 2  
 US-09-821-687-4 (1-561) x AX011753 (1-3101)  
 Qy 50 G1yLeuValAlaH1eSerPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 69  
 Db 89 GGGCTAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 148  
 Qy 70 G1uApsG1yAlaLeuAlaValLeuGlnGlnIlePheLysSerPaaPaaPaaPaaPaaPaaPaa 89  
 Db 149 GAGAG 208  
 Qy 90 AsnLysSerAlaPheLeuCysG1yValMetLysThrTyrArgG1uArgG1uLysGlnGln 109  
 Db 209 AACAAAAGGCGCTTTTATAGTGAAGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 268  
 Qy 110 ThrLysValAlaAspSerSerLysG1yProPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 129  
 Db 269 ACCAAGAG 328  
 Qy 130 ArgThrG1yTyrThrLeuAspValThnThnG1yGlnArgG1yG1yLysProProPro 149  
 Db 329 AGAAGAGGCTACCACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 388  
 Qy 150 AspSerValTyrSerG1yGlnGlnProSerValG1yThnG1uIlePheValG1yLysIle 169  
 Db 389 GATTCGTTTATTAAGGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 448

QY 170 ProArgAspLeuPheGluAspGluLeuValProLeuPheGluValGlyProIleTyr 189  
 Db 449 CCAAGAAATCTATTGAGGAGGAACTGTTTCATTATTGAGAAAGCTGAGCTTATTTGG 508  
 QY 190 AspLeuArgLeuMetMetAspProLeuThrGlyLeuAsnArgGlyTyrAlaPheValThr 209  
 Db 509 GATTTGCTCTAATGATGAGTCACTGCTGCTCAATAGAGTTATGCTTTGCTACT 568  
 QY 210 PheCysThrLygGluAlaAlaGlnGluAlaVallyLeuTyrAsnAsnHiseGlnIleArg 229  
 Db 569 TTTTGTCAAAAAGAGACGCTCAGAGGCTGCTTAACGTATTAATCAATGAAATTCGT 628  
 QY 230 SerGlyLyshHisIleGlyValIleSerValAlaAsnAsnArgLeuPheValGlySer 249  
 Db 629 TCTGAAAACATATTTGGTGTGCTGCTCACTGCTCAACATAGGCTTTTGTGGGCTCT 688  
 QY 250 IleProLySerLyshThrLygGluGlnIleLeuGluGluPheSerLyshValThrGluGly 269  
 Db 689 ATTCTTAAGACTTAACCAAGAACAGATTTCTTGAAGAAATTACCAAGTAAACAGAGGCT 748  
 QY 270 LeuThrAspValIleLeuTyrHisGlnProAspArgLyshLyshAsnArgGlyPheCys 289  
 Db 749 CTTCACACGCTCATTTTATACCAACACCGGATGACAAAGAAAAACAGAGGCTTTTGC 808  
 QY 290 PheLeuGluTyrGluAspHisIleThrAlaGlnIleAlaArgArgLeuMetSerGly 309  
 Db 809 TTTCTTGAATATGAGATCAACAAACAGCTGCCAG----- 844  
 QY 310 LyshVallyshValTyrGlyAsnValGlyThrValGluTyrAlaAspProIleGluAspPro 329  
 Db 844 ----- 844  
 QY 330 AspProGluValMetAlaLyshVallyshValLeuPheValAlaGlnLeuAlaAsnThrVal 349  
 Db 845 -----GTAAGAGTCTGCTTTGTGACCAACCTTGCCAAATACGTGA 883  
 QY 350 ThrGluGluIleLeuGluLyshSerPheSerGlnPheGlyLyshLeuGluArgVallysh 369  
 Db 884 ACAGAAAGATTTTGAAGAAAGCACTTATGACGTTTGGAAACGTGAACAGATGAAGAG 943  
 QY 370 LeuLyshAspTyrAlaPheIleHisPheAspGluArgAspGlyAlaVallyshAlaMetGlu 389  
 Db 944 TTTAAAGATTATGGCTTCACTTATGATGAGGAGATGAGTGTCTCAACGCTATGAA 1003  
 QY 390 GlnMetAsnGlyLyshAspLeuGluGlyGluAsnIleGluIleValPheAlaLyshProPro 409  
 Db 1004 GAAATGATGCAAAAGACTTGAGAGGAGAAATATTTGAAATTGTTTGCACAACTCA 1063  
 QY 410 AspGlnLyshArgLyshGluArgLyshAlaGlnArgGlnAlaAlaLyshAsnGlnMetTyrAsp 429  
 Db 1064 GATGAGAAAGAGAAAGAAAGAAAGCTCAGAGGACAGCAAGCAAAATCAATGATATGAC 1123  
 QY 430 AspTyrTyrTyrTyrGlyProProHisMetProProProThrArgGlyArgGly 449  
 Db 1124 GATTACTACTATATGCTCCTCATATGCTCCCTCAACAAGAGCTCGAGGGGCTGGA 1183  
 QY 450 GlnArgGlyGlyTyrGlyTyrProProAspTyrTyrGlyTyrGluAspTyrTyrAspTyr 469  
 Db 1184 GGTGAGAGGTGATATGATATCTCCAGATTAATATGATATGAAATTAATATGATATAT 1243  
 QY 470 TyrGlyTyrAspTyrHisAsnTyrArgGlyGlyTyrGluAspProTyrTyrGlyTyrGlu 489  
 Db 1244 TATGATATATGATATCAATCAATCAATGCTGATATGAAATCAATCAATCAATGATATGAA 1303  
 QY 490 AspPheGlnValGlyAlaArgGlyArgGlyGlyArgGlyAlaArgGlyAlaAlaProSer 509  
 Db 1304 GATTTTCAAGTTGAGAGCTAGAGAGAGGGCTGTAGAGAGCAAGGGGCTGCTCCATCC 1363  
 QY 510 ArgGlyArgGlyAlaAlaProProArgGlyArgGlyAlaGlyTyrSerGlnArgGlyGlyPro 529  
 Db 1364 AGAGGTCTGGGGCTGCTCTCCCTCCCGGTAGAGCGGCTATTCACAGAGAGAGGTCTCT 1423

QY 530 GlySerAlaArgGlyValArgGlyAlaArgGlyValGlnGlnGlnArgGlyArgGly 549  
 Db 1424 GATTCAGCAAGAGCGCTTCAGGCTGCGAGAGAGGCGCCCAACAAAGAGCGCGG 1483  
 QY 550 ---GlyArgGlyValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 561  
 Db 1484 CAGGAAAGAGGCTCAGAGCGGCTCTGACCTTTACAA 1522  
 RESULT 14  
 AF441128  
 LOCUS  
 DEFINITION Mus musculus CD-1 heterogeneous nuclear ribonucleoprotein R mRNA,  
 complete cds.  
 ACCESSION AF441128  
 VERSION AF441128  
 KEYWORDS AF441128.1 GI:17066598  
 SOURCE  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1899)  
 Rosolli, W., Kroning, A.K., Ohndorf, U.M., Steegborn, C., Jablonka, S.  
 and Sendtner, M.  
 TITLE Specific interaction of Smn, the spinal muscular atrophy  
 determining gene product, with hnRNP-R and gryn-1/hnRNP-Q: a role  
 for Smn in RNA processing in motor axons?  
 Hum. Mol. Genet. 11 (1), 93-105 (2002)  
 MEDLINE 21634625  
 PUBMED 11773003  
 REFERENCE 2 (bases 1 to 1899)  
 Rosolli, W., Kroning, A.-K. and Sendtner, M.  
 AUTHORS Direct Submission  
 TITLES Submitted (30-OCT-2001) Inst. f. Clinical Neurobiology, Wuerzburg  
 University, Josef-Schneider Str. 11, Wuerzburg 97080, Germany  
 FEATURES  
 source location/Qualifiers  
 1..1899  
 /organism="Mus musculus"  
 /strain="CD-1"  
 /db\_xref="taxon:10090"  
 1..1899  
 /note="hnRNP-R"  
 /codon\_start=1  
 /product="heterogeneous nuclear ribonucleoprotein R"  
 /protein\_id="AA13532.1"  
 /db\_xref="GI:17066598"  
 /translation="MANVNGNAVOLKEEPEMDTSSVTHTEHYTLLEAGLPQYAE  
 RIDEIFQTELVAVYVDLDERALIDALREPEBEGALSVLOQKESDLRHVOKSAFLGVN  
 KTYROREKSGSVQESTKGPDEAKIKALERTGTLVDYTGKRTGPPDPDVGVO  
 PGIGTEVFVGIKPIVDLDEDLVLFEEKAPVLDLRMDPLSGORGVAFITFGKEA  
 AGAVKICDSVETIRPGKHLGVCISVANNRLFGSTPKNKTENITIEEPEKTEGELVDV  
 ILYHOPDDKKRGRGCTLYEPDHKSAQAQRRLMGSKYKVMKNVTVTEADVEEPDP  
 EVMAKTVLFRNLATVTYETLETLESFSEFGLEVKIKDYAFVFEDEGAVAVAMD  
 ENNGKEIEBERIEIVLAKPPDKRKRRKROARQASRSYAEVDYHPPPMPPMGRG  
 RGRGGYGYPPDYGYEDYDDYGYDYHDYDGVEDPYGDDGAYAVAGRGGRGGR  
 GAPPGRGGRGPRGAYSGRGAIPGPRGSRGGRGGRGGRGGRGGRGGRGGRGGRGGRG  
 NVGGRKXKADGVNQPDSKRGQTNNQGNWSQPIAQGPLQGGGYSNRYGNNDNQERYO  
 DRYGQMK"  
 BASE COUNT 581 a 362 c 548 g 408 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7, 07e-105 Length: 1899  
 Score: 2413.50 Matches: 458  
 Percent Similarity: 90.39% Conserved: 50  
 Best Local Similarity: 81.49% Mismatches: 41  
 Query Match: 80.94% Indels: 13  
 DB: 10 Gaps: 9  
 US-09-821-687-4 (1-561) x AF441128 (1-1899)  
 QY 1 MetAlaThrGluHisValaGlnGlyArgGlyThr-----GluGluProMet 15  
 Db 1 ATGGCTAATCAG---GTGAATGTTAATGGCGTACGTTAAAGAGAGGAAACCAATG 57

16 AspThrThrsAlaValIleHisSerGluAsnPhenGlnThrLeuLeuAspAlaGlyLeu 35  
 58 GAT---ACTTCCAGTGTACTACACAGAACATCAACAGACCTGATAGAGGAGGCTTC 114  
 36 ProGlnLysValAlaGlnLysLeuAspGlnIleTyrValAlaGlyLeuValAlaHisSer 55  
 115 CCACAGAGGCTGCAGAAAGACCTTGAATGATATTTTCAGACAGATTTGGTATGATG 174  
 56 AspLeuAsnGlnLysValAlaGlnLysLeuLeuGlnLysAsnGlnLysAlaLeuAla 75  
 175 GATCTGATATAAGAGCAATCGATGCTCTCAGAGGATTTATAGAGAGAGAGCTGTCT 234  
 76 ValLeuGlnLysPheLysAspSerAspLeuSerHisValAlaGlnLysSerAlaPheLeu 95  
 235 GTATCGAGAGCTTCAAGAAAGATGACCTGTCTCATCTCCAGAACAAATGCTATTTTA 294  
 96 CysGlyValMetLysThrTyrArgGlnArgGlnLysGlnLysThrValAlaAspSer 115  
 295 TGTGAGATTATGAAGACCTACAGGCAAGAGAAACAGGACAGCAAGTACAAAGAGTCC 354  
 116 SerLysGlyProAspGlnLysValAlaGlnLysLeuLeuGlnLysThrGlyTyrThrLeu 135  
 355 ACCAAAGGCGCCATGAGAAATCAAGGCTTGTGAGAGAGCTGCTATTAATTG 414  
 136 AspValThrThrGlyGlnArgLysTyrGlyProProAspSerValTyrSerGly 155  
 415 GATGTACACCGGAGCAGAGAGATAGCGGCTCCCAACAGACGTGTGACTCGAGT 474  
 156 GlnGlnProSerValGlyThrGlnLysPheValGlyLysIleProArgAspLeuPheGln 175  
 475 GTCCACCTGTGAGTGGAGCAGAGGCTTCTGTAGTAAATCCCAAGACCTGTATGAG 534  
 176 ArgGlnLeuValProLeuPheGlnLysAlaGlyProIleThrAspLeuGlnMetMet 195  
 535 GATGAATATAGTCCCTTTTGAAGAGCTGTGCTATCGGAGTTCACGCTCATGATG 594  
 196 AspProLeuThrGlyLeuAsnArgGlyTyrValPheValThrPheCysThrLysGlnAla 215  
 595 GATCCACTCTCTGTGAGAGCAGAGGCTATGCTATTCATCCTTGTGAAAGAGAGCT 654  
 216 AlaGlnGlnLysValLysLeuTyrAsnAsnHisGlnIleArgSerGlyLysHisIleGly 235  
 655 GCGCAGAGACAGTAAAGCTGTGTGACAGTGAATGATGCTCCCTGTAAACCTGGGA 714  
 236 ValCysIleSerValAlaAsnAsnArgLeuPheValGlySerIleProLysSerLysThr 255  
 715 GTGTGATTTCTGTGCAACCAACAGCTTTTGTGTGATCCATTCACAAAGAAATAAACT 774  
 256 LysGlnGlnIleLeuGlnLysPheSerLysValThrGlnLysLeuThrAspValIleLeu 275  
 775 AAGAGAAACCTTGTGAGAGGCTTCAAGTCAACAGAGGCTTTGGTGCAGCTTATTTCT 834  
 276 TyrHisGlnProAspArgLysLysLysAsnArgGlyPheCysPheLeuGlnLysAsp 295  
 835 TATCATCAACCCGATGCAAAAGAGAGATCGGGGCTTCTGCTTCAATATAGAGAT 894  
 296 HisLysThrAlaAlaGlnLysArgGlyArgLeuMetSerGlyLysValLysValTyrGly 315  
 895 CACAAGTCAGACACCAAGCAGAGCGGCTGATGAGTGAAGAAAGTCAAGAGCTGGGGA 954  
 316 AsnValGlyThrValGlnTyrAlaAspProIleGlnLysProAspProGlnValMetAla 335  
 955 AATGTGTTCAGTAAAGTGGCTGACCCCGTGAAGAGCGGATCCAGAGTCAAGTGGCA 1014  
 336 LysValLysValLeuPheValArgAsnLysValAsnThrValThrGlnLysIleLeuGln 355  
 1015 AAGGTAAAGATTATTTGTGAGAACTTGGCACTACGATGACAGAGAAATTTAGAG 1074  
 356 LysSerPheSerGlnPheGlyLysLeuGlnLysValLysLysLeuLysAspTyrAlaPhe 375  
 1075 AAGTCAATTTCTAGATTGAGAAACCTGAGCGAGTGAAGAAAGTTGAGATTAAGCTTC 1134

376 IleHisPheAspGlnLysArgAspGlyAlaValLysAlaMetGlnLysMetAsnGlyLysAsp 395  
 1135 GTTCAATTTGAGACACGGGAGCTGCTTTAGCTATAGATGAAATGATGAAAGAA 1194  
 396 LeuGlnGlyLysAsnIleGlnLysValPheAlaLysProProAspGlnLysArgLysGln 415  
 1195 ATGAAAGGGGAAAGAAATTTGATTTGCTTAACTTACCTACCAACAGACAAAGAAAGAG 1254  
 416 ArgLysAlaGlnLysGlnLysAlaLysAsnGlnMetTyrAspAspTyrTyrTyrGly 435  
 1255 CGCAGCTGTAGACAGGCTCCAGAACACTGCGTATGAAGATTTATATATACCCCT 1314  
 436 ProProHisMetProProProThrArgGlyArgGlyArgGlyValArgGlyLysGly 455  
 1315 CTTCCCGGATGCTCACTTCAATGAGAGCCCAAGGCTGTGAGAGAGAGTGAATGAGC 1374  
 456 TyrProProAspTyrTyrGlyLysAspTyrTyr---AspTyrTyrGlyTyrAspTyr 474  
 1375 TACCTCCAGATTTATATGCTATGAGAACTATGATGATGATTAATGATTTAT 1434  
 475 HisAsnTyrArgGlyLysTyrGlnAspProTyrTyrGlyTyrGlnLys---PheGlnVal 493  
 1435 CACGACTATCGTGAAGCTATGAAAGATCCCTACTACGCTATGATGATGATGATGACGTA 1494  
 494 GlyAlaArgGlyValArgGlyValArgGlyAlaArgGlyAla---AlaProSerArgGlyArg 512  
 1495 AGAGAGAGAGCA---GAGAGAGAGGAGGAGGAGGAGGCTTCCACACACAGAGGAGGCG 1551  
 513 GlyAlaAlaProProArgGlyValArgAlaGlyTyrSerGlnArgGlyLysPro---GlySer 531  
 1552 GAGGACACACCTCCAGAGAGTATGAGCTGCTATTCACAGAGGAGGACCTTTGGAGCA 1611  
 532 AlaArgGlyValArgGlyAlaArgGlyGly---AlaGlnGlnLysArgGlyValArgGly 550  
 1612 CCAAGAGGCTCAAGAGGCTGAGAGGAGGCTCTGCAACAGACAGAGAGCGGTGTTCT 1671  
 551 LysGly 552  
 1672 CTTGGA 1677

RESULT 15  
 AF000364  
 LOCUS  
 DEFINITION  
 Home sapiens heterogeneous nuclear ribonucleoprotein R mRNA,  
 complete cds.  
 ACCESSION  
 AF000364  
 VERSION  
 AF000364.1 GI:2697102  
 SOURCE  
 ORGANISM  
 Homo sapiens.  
 Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 (bases 1 to 2644)  
 Haefield,W., Chan,B.K., Mathison,D.A., Portman,D., Dreyfuss,G.,  
 Steiner,G. and Tan,E.M.  
 Molecular definition of heterogeneous nuclear ribonucleoprotein R  
 (hnRNP R) using autoimmune antibody: immunological relationship  
 with hnRNP P  
 Nucleic Acids Res. 26 (2), 439-445 (1998)  
 JOURNAL  
 MEDLINE  
 PUBMED  
 9421497  
 98083170  
 2 (bases 1 to 2644)  
 Haefield,W., Chan,B.K. and Tan,E.M.  
 Direct Submision  
 Submitted (18-Apr-1997) Molecular and Experimental Medicine, The  
 Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla,  
 CA 92037, USA  
 FEATURES  
 source  
 1..2644  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="BL14"  
 /cell\_line="T24"



Mon Apr 28 06:48:47 2003

Search completed: April 28, 2003, 06:50:02  
Job time : 8634 secs

us-09-821-687-4.rge

Page 21

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4 ps 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2003, 16:49:41 ; Search time 17371.6 Seconds

(without alignments)  
5783.170 Million cell updates/sec

Title: US-09-821-687-3

Perfect score: 3452  
Sequence: 1 gcgggttcgtcgcagcgsc.....ggatcccggtgacgagct 3452

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenBank1.\*  
1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_ph.\*  
7: gb\_pl.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_ro.\*  
11: gb\_scs.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_ov.\*  
22: em\_or.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_scs.\*  
28: em\_un.\*  
29: em\_vl.\*  
30: em\_hcg\_hum.\*  
31: em\_hcg\_inv.\*  
32: em\_hcg\_other.\*  
33: em\_hcg\_mus.\*  
34: em\_hcg\_pln.\*  
35: em\_hcg\_rod.\*  
36: em\_hcg\_mam.\*  
37: em\_hcg\_vrt.\*  
38: em\_sy.\*  
39: em\_hlgo\_hum.\*  
40: em\_hlgo\_mus.\*  
41: em\_hlgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3452	100.0	3452	10	AB035725
2	2741.4	79.4	3319	9	BC032643
3	2542.2	73.6	3101	6	AX011753
4	1953.8	56.6	2221	9	AF155568
5	1886.6	54.7	2195	10	BC004001
6	1826.2	52.9	2079	6	AR177851
7	1802	52.2	103819	9	HS1775C13
8	1802	52.2	212659	2	AL1390737
9	1674	48.5	2932	9	AF037448
10	1608	46.6	2373	10	AF093821
11	1584.8	45.9	1686	9	AY034483
12	1552.2	45.0	1872	9	AY034481
13	1456.8	42.2	2232	10	AF408434
14	1356.4	39.3	74337	9	AL136082
15	1343.4	38.9	1767	9	AY034482
16	1156	33.5	221309	2	AC122217
17	1033	29.9	175921	2	AC111832
18	957.6	27.7	151297	2	AC118904
19	896	26.0	1899	10	AF441128
20	875	25.3	2644	9	AF000364
21	856	24.8	2722	9	BC001449
22	853.8	24.7	2471	10	BC004679
23	816	23.6	181810	2	AC126148
24	762.4	22.1	121304	2	AC079134
25	762.4	22.1	196488	2	AC084410
26	742.2	21.5	202374	10	AC122868
27	736.4	21.3	934	6	AX013715
28	636.8	18.4	139996	9	AL161799
29	570.6	16.5	129237	2	AC127932
30	484	14.0	110783	2	AC129693
31	457.2	13.2	211030	2	AC125224
32	426	12.3	118429	2	AC119124
33	398.4	11.5	57824	2	AC101053
34	381.8	11.1	2606	3	AY058477
35	368	10.7	384	6	AR079026
36	309.2	9.0	413	9	AR079027
37	289.4	8.4	2467	9	BC028588
38	250.8	7.3	1785	10	AY028945
39	250.8	7.3	1863	10	AF290984
40	247	7.2	175221	2	AC111832
41	242.4	7.0	1770	9	HS4272079
42	242.4	7.0	1794	9	HS4272078
43	242.4	7.0	1945	9	AF209192
44	240.8	7.0	1761	9	AF271789
45	233	6.7	1911	9	AF271790

# ALIGNMENTS

RESULT 1  
AB035725  
LOCUS  
DEFINITION Mus musculus SYNCRIP mRNA, complete cds.  
ACCESSION AB035725  
VERSION AB035725.1 GI:6576814  
KEYWORDS SYNCRIP.  
SOURCE Mus musculus  
ORGANISM Mus musculus (strain:DDV) postnatal 21 day CDNA to mRNA.  
REFERENCE  
AUTHORS Mizutani,A., Fukuda,M., Ibatu,K., Shiratschi,Y. and Mikoshiba,K.  
TITLE 1 (sites)  
SYNCRIP, a cytoplasmic counterpart of heterogeneous nuclear ribonucleoprotein R, interacts with ubiquitous synaptotagmin

isoforms  
JOURNAL J. Biol. Chem. 275 (13), 9823-9831 (2000)  
MEDLINE 20200483  
REFERENCE 2 (bases 1 to 3452)  
AUTHORS Mizutani, A., Fukuda, M. and Mikoshiba, K.  
TITLE Direct Submission  
JOURNAL Submitted (11-DEC-1999) Akihito Mizutani, Brain Science Institute,  
RIKEN, the Laboratory for Developmental Neurobiology, 2-1 Hirosawa,  
Mako, Saitama 351-0198, Japan (E-mail: amizutani@ims.u-tokyo.ac.jp,  
Tel: +81-48-467-9745, Fax: +81-48-467-9744)

## FEATURES

source  
1. .3452  
/organism="Mus musculus"  
/strain="DDY"  
/db\_xref="taxon:10090"  
/dev\_stage="postnatal 21 day"  
1. .3452  
/gene="SYNCRIP"  
1. .3452  
/gene="SYNCRIP"  
/evidence="experimental"  
154. .1839  
/gene="SYNCRIP"  
/codon\_start=1  
/protein\_id="BAA8342.1"  
/db\_xref="GI:6576815"  
/translation="MATEHYNGNGTEBPMDTTSAVIHSNPTLLDAGLPOKAEFLD  
EIVYAGLVNHSDDLDERAIEALKKPNEDGALAVIQKRDLSLHVOKSAFLCGMKEY  
ROREKQGTIVADSSKDPDEAKIPLALERTGYTLDTVTGQKRGSPVSYSGQPSV  
GTEIFVQKIPRDLPEDELVPLPEKAPIMDLRLMDPLGLNNGVAFVSTGEAIOE  
AVLNNHRIIRSGKHIQVCSIVANNRLFVSIKSKTEQILIEFSEKTEGLDVLV  
HOPDDKKRGRFCFLEBEDHKTAAORRLMSGVNMGVGVENADPIBDDEPVM  
AKKVLFTNLTATTEBILIEKSFSGKGLERKGLKDYAFHEDRDVAQWEMEM  
GDLSEBNEIIVAKPPDQKRRKRAQRAKQMDTDYTYGPPEPPTTRKRGEG  
KGGTGPDTGYGIEDYTDYDYNHRYRGSTGYEDPYGDFQVAGRGGRGGRGAP  
SRGRGAAPPGRGYSQRGPGSGARGVARGAGAOORGRGKGVABDLO"

BASE COUNT 1056 a 623 c 777 g 996 t  
ORIGIN

Query Match 100.0%; Score 3452; DB 10; Length 3452;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGGTTTGTCCAGAGCGGCGTTCGGCGCCATTATTAACAAGCTCCACTCGCGC 60  
DB 1 GCGGGTTTGTCCAGAGCGGCGTTCGGCGCCATTATTAACAAGCTCCACTCGCGC 60  
QY 61 CGGACACAGGAGAGCAGGAGCAGCGGTTCCCGCAGCCGACCCGATCGAAGATTCT 120  
DB 61 CGGACACAGGAGAGCAGGAGCAGCGGTTCCCGCAGCCGACCCGATCGAAGATTCT 120  
QY 121 CGGCGCCAGCGCCCGCGGGGAGATCTCTGGAACAATGGCTACAGAACTGTTAAAGAAAT 180  
DB 121 CGGCGCCAGCGCCCGCGGGGAGATCTCTGGAACAATGGCTACAGAACTGTTAAAGAAAT 180  
QY 181 GGTACTGAAGAGCCCATGATCTACTTCAAGCACTTATCCATTGAGAAAATTTTCAGACA 240  
DB 181 GGTACTGAAGAGCCCATGATCTACTTCAAGCACTTATCCATTGAGAAAATTTTCAGACA 240  
QY 241 TTGCTTATCTGCTGTTTACCAAGAAAGTTGCTGAAAACCTAGATGAATTTACGTTGCA 300  
DB 241 TTGCTTATCTGCTGTTTACCAAGAAAGTTGCTGAAAACCTAGATGAATTTACGTTGCA 300  
QY 301 GGGCTAGTTCACATATGATTTTATGAGAAAGACTATCGAAGCTTTTAAAGAGTTCAAT 360  
DB 301 GGGCTAGTTCACATATGATTTTATGAGAAAGACTATCGAAGCTTTTAAAGAGTTCAAT 360  
QY 361 GAAGACGGCGCATTTGGAGTCTTCAAGATTAAAGCAGTATCTCTCATGTTTCAAG 420  
DB 361 GAAGACGGCGCATTTGGAGTCTTCAAGATTAAAGCAGTATCTCTCATGTTTCAAG 420  
QY 421 AACAAAAGTCCCTTTTATGTGAGTCAAGAGCTTACAGCGCAGAGAAAACAGGGG 480  
DB 421 AACAAAAGTCCCTTTTATGTGAGTCAAGAGCTTACAGCGCAGAGAAAACAGGGG 480

DB 421 AACAAAAGTCCCTTTTATGTGAGTCAAGAGCTTACAGCGCAGAGAAAACAGGGG 480  
QY 481 ACCAAAGTAGACACTCTAGTAAAGACCAAGATGAGGCAAAATTAAGCACTTTGGAA 540  
DB 481 ACCAAAGTAGACACTCTAGTAAAGACCAAGATGAGGCAAAATTAAGCACTTTGGAA 540  
QY 541 AGAAGAGGTACACACTGTGATGACATCAAGGTCAGAGGAAAGTATGAGAGCACTTCCA 600  
DB 541 AGAAGAGGTACACACTGTGATGACATCAAGGTCAGAGGAAAGTATGAGAGCACTTCCA 600  
QY 601 GATTCGGTTTATTCAGATCAGACGCTTCTGTGGCACTGAGATATTTTGGGAAAGATC 660  
DB 601 GATTCGGTTTATTCAGATCAGACGCTTCTGTGGCACTGAGATATTTTGGGAAAGATC 660  
QY 661 CCAGAGATCTGTTTGAAGATGAGCTTGTTCCTATTATTTGAAGAACTGACCTTATATGG 720  
DB 661 CCAGAGATCTGTTTGAAGATGAGCTTGTTCCTATTATTTGAAGAACTGACCTTATATGG 720  
QY 721 GATCTTGTATTAAGATGATCCGCTCACTGATCTCAACAGAGATTATGCTTGTCACT 780  
DB 721 GATCTTGTATTAAGATGATCCGCTCACTGATCTCAACAGAGATTATGCTTGTCACT 780  
QY 781 TTTTGTACAAAAGACAGACAGACAGGCTGTAACTGTATATATATCATGAAATTCGT 840  
DB 781 TTTTGTACAAAAGACAGACAGACAGGCTGTAACTGTATATATATCATGAAATTCGT 840  
QY 841 TCCGGAGAGACATTTGCTGTCTGATCTCACTGATGCAATATGCTTTTGTGGGCTCG 900  
DB 841 TCCGGAGAGACATTTGCTGTCTGATCTCACTGATGCAATATGCTTTTGTGGGCTCG 900  
QY 901 ATTCTTAAGATTAACCAAGAGCAGATTCCTGAAGATTTAGCAAGATGACAGAGGGT 960  
DB 901 ATTCTTAAGATTAACCAAGAGCAGATTCCTGAAGATTTAGCAAGATGACAGAGGGT 960  
QY 961 CTCACAGATGTCATTTTATACCAACCACTGATGACAGAAATTAAGAGCTTTTGG 1020  
DB 961 CTCACAGATGTCATTTTATACCAACCACTGATGACAGAAATTAAGAGCTTTTGG 1020  
QY 1021 TTTCTGATATGAAATGATCACAAAACAGCTGCCAGGCAAGCTTATGATGATGCT 1080  
DB 1021 TTTCTGATATGAAATGATCACAAAACAGCTGCCAGGCAAGCTTATGATGATGCT 1080  
QY 1081 AAGTCAAAAGTCGGGAAATGTTGGAACGTTGAGTGGCTGATCTTATGAAATTCCT 1140  
DB 1081 AAGTCAAAAGTCGGGAAATGTTGGAACGTTGAGTGGCTGATCTTATGAAATTCCT 1140  
QY 1141 GATCTGAAGTATGCAAAAGGTAAGTGTCTTTGTACGCAACCTTGCACACAGGTA 1200  
DB 1141 GATCTGAAGTATGCAAAAGGTAAGTGTCTTTGTACGCAACCTTGCACACAGGTA 1200  
QY 1201 ACAGAAAGAAATTTTAAAGATGATTTAGCTTTGGGAACTGGAAGAGTAAAGAAAG 1260  
DB 1201 ACAGAAAGAAATTTTAAAGATGATTTAGCTTTGGGAACTGGAAGAGTAAAGAAAG 1260  
QY 1261 CTAAGAGTATGCTTTCATTCATTTTGAAGAGATGCTGCTCAAGGCTTATGGA 1320  
DB 1261 CTAAGAGTATGCTTTCATTCATTTTGAAGAGATGCTGCTCAAGGCTTATGGA 1320  
QY 1321 GAAATGAATGTAAGACTTGAAGGAGAAATATTTGAAATGTTTTGCTAAGCCACA 1380  
DB 1321 GAAATGAATGTAAGACTTGAAGGAGAAATATTTGAAATGTTTTGCTAAGCCACA 1380  
QY 1381 GATCAGAAAGGAAAGAAAGAAAGCTCAGAGGCAAGCAGCAAGATCAATGATGAT 1440  
DB 1381 GATCAGAAAGGAAAGAAAGAAAGCTCAGAGGCAAGCAGCAAGATCAATGATGAT 1440  
QY 1441 GATTACTAATATATGCTCACTCATATATGCTCCCAACAGAGGTGAGGGGCTGGA 1500  
DB 1441 GATTACTAATATATGCTCACTCATATATGCTCCCAACAGAGGTGAGGGGCTGGA 1500  
QY 1501 GGTAGAGTGTGCTATGATATCTCCAGATTTATGATGATGATGATGATGAT 1560  
DB 1501 GGTAGAGTGTGCTATGATATCTCCAGATTTATGATGATGATGATGATGAT 1560

QY 1561 TATGTTATGATTACCTAATCTATCGTGTGATATGAAGATCCATCTATGTTATGAA 1620  
 DB 1561 TATGTTATGATTACCTAATCTATCGTGTGATATGAAGATCCATCTATGTTATGAA 1620  
 QY 1621 GATTTTCAAGTTGAGCTAGAGAGAGGGGTGTAGAGAGCAAGGGGTGTCTCCATCC 1680  
 DB 1621 GATTTTCAAGTTGAGCTAGAGAGAGGGGTGTAGAGAGCAAGGGGTGTCTCCATCC 1680  
 QY 1681 AGAGGTGTGAGGGGTGTCTCTCCCGGTGTAGAGCCGGTTATTCAGAGAGAGGGCCCT 1740  
 DB 1681 AGAGGTGTGAGGGGTGTCTCTCCCGGTGTAGAGCCGGTTATTCAGAGAGAGGGCCCT 1740  
 QY 1741 GATTCAGAGAGAGGGGTGTCTCTCCCGGTGTAGAGAGGGGTGTCTCCATCC 1800  
 DB 1741 GATTCAGAGAGAGGGGTGTCTCTCCCGGTGTAGAGAGGGGTGTCTCCATCC 1800  
 QY 1801 GGAAGAGGGGTGTAGAGCCGGTGTCTGACCTGTTCATGAAGAGCTGATTTGTGGG 1860  
 DB 1801 GGAAGAGGGGTGTAGAGCCGGTGTCTGACCTGTTCATGAAGAGCTGATTTGTGGG 1860  
 QY 1861 ATTAACCAAGAGCTTGCATGTGATGATATGTAAGAGAAATCAAGCACTTAAATCT 1920  
 DB 1861 ATTAACCAAGAGCTTGCATGTGATGATATGTAAGAGAAATCAAGCACTTAAATCT 1920  
 QY 1921 CGATGTATGAGAGCATATCTTATGCAAGAGCCCTCTATGAGATCANGAATGAAA 1980  
 DB 1921 CGATGTATGAGAGCATATCTTATGCAAGAGCCCTCTATGAGATCANGAATGAAA 1980  
 QY 1981 TACGGAGCATGTGAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 2040  
 DB 1981 TACGGAGCATGTGAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 2040  
 QY 2041 AGTGCAGATTTAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2100  
 DB 2041 AGTGCAGATTTAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2100  
 QY 2101 ATAACTTCTAGAGCATATCTTATGCAAGAGCCCTCTATGAGATCANGAATGAAA 2160  
 DB 2101 ATAACTTCTAGAGCATATCTTATGCAAGAGCCCTCTATGAGATCANGAATGAAA 2160  
 QY 2161 TAAAGATTTTCCCTCTTTTATTAAGATCTCAGAGTGTATCTGAAATAT 2220  
 DB 2161 TAAAGATTTTCCCTCTTTTATTAAGATCTCAGAGTGTATCTGAAATAT 2220  
 QY 2221 CTCTGATTTTGGCATTTAAATTTTATGCAAGTATCAGAGTGTATCTGAAATAT 2280  
 DB 2221 CTCTGATTTTGGCATTTAAATTTTATGCAAGTATCAGAGTGTATCTGAAATAT 2280  
 QY 2281 TAAATTTTATTTCTGCTCAAGAGTAAATGATGATGATGATGATGATGATGATGAT 2340  
 DB 2281 TAAATTTTATTTCTGCTCAAGAGTAAATGATGATGATGATGATGATGATGATGAT 2340  
 QY 2341 TCACTTTTGAATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400  
 DB 2341 TCACTTTTGAATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400  
 QY 2401 GCATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460  
 DB 2401 GCATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460  
 QY 2461 GCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520  
 DB 2461 GCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520  
 QY 2521 AAGGCTTGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580  
 DB 2521 AAGGCTTGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580  
 QY 2581 TTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640  
 DB 2581 TTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640

QY 2641 TTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700  
 DB 2641 TTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700  
 QY 2701 TGTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760  
 DB 2701 TGTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760  
 QY 2761 CATCCGATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820  
 DB 2761 CATCCGATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820  
 QY 2821 AATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2880  
 DB 2821 AATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2880  
 QY 2881 TAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940  
 DB 2881 TAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940  
 QY 2941 ATTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000  
 DB 2941 ATTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000  
 QY 3001 TTGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060  
 DB 3001 TTGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060  
 QY 3061 CATGAT 3120  
 DB 3061 CATGAT 3120  
 QY 3121 AAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180  
 DB 3121 AAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180  
 QY 3181 TAAAGAAACACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 3240  
 DB 3181 TAAAGAAACACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 3240  
 QY 3241 GCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300  
 DB 3241 GCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300  
 QY 3301 TAAATGCAATTTTAAACAGAAATTTTAAACAGATGATGATGATGATGATGATGATGATGAT 3360  
 DB 3301 TAAATGCAATTTTAAACAGAAATTTTAAACAGATGATGATGATGATGATGATGATGATGAT 3360  
 QY 3361 AATCATTTTGCATTTCTGCAATTTTAAATTTTAAACAGATGATGATGATGATGATGATGATGAT 3420  
 DB 3361 AATCATTTTGCATTTCTGCAATTTTAAATTTTAAACAGATGATGATGATGATGATGATGATGAT 3420  
 QY 3421 AACTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480  
 DB 3421 AACTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480

RESULT 2  
 BC032643  
 LOCUS 3319 bp mRNA linear PRI 27-JUN-2002  
 DEFINITION Homo sapiens, similar to NS1-associated protein 1, clone MGC:45213  
 IMAGE:5495201, mRNA, complete cds.  
 ACCESSION BC032643  
 VERSION BC032643.1 GI:21619167  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 3319)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian

REMARK  
COMMENT

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@hgrl.nih.gov](mailto:nisc_mgc@hgrl.nih.gov)  
 Akhter N., Ayele K., Beckstrom-Sternberg S.M., Benjamin B., Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S., Dietrich N.U., Granite S., Guan X., Gupta J., Hashiguchi P., Hansen N., Ho S.-L., Karlins E., Kwong P., Latic P., Legaapi R., Maduro Q.L., Mastello C., Maskeri B., Mastrian S.D., McCloskey J.C., McDowell J., Pearson R., Stantipop S., Thomas P.J., Touchman J.W., Teuregon C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L., Young A., Zhang L.-H. and Green E.D.

## FEATURES

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLNL at: <http://image.lnl.gov>  
 Series: IRAC Plate: 69 Row: f Column: 23  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5453805.  
 Location/Qualifiers

1..3319

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:45213 IMAGE:5495201"  
 /issue\_type="Eye, retinoblastoma"  
 /clone\_lib="NIH\_MGC\_67"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORT6"  
 98..1681  
 /codon\_start=1  
 /product="Similar to NS1-associated protein 1"  
 /protein\_id="AAH32643.1"  
 /db\_xref="gi:21619168"

## CDS

/translation="MATEHVNNGTEBPMDDTSVAVHSENFQTLIDAGLPQVAKELD EIVAGLVASHSDERAIKALKEFNEDGALAVQKPSDLSHONKSAFLCVMKTY ROBEKGTKVADSKSPDEAKIKALERTYGLTDVTRGKRGKPPDVSQOQSV GREIFVGI PRDIFDEIVPLPERAGIPMDLRIMDPLRGKRGKPPDVSQOQSV AKLVNNHETRSGKHICVCTSVANNRLFGSITKSTKTEQILIEFSKVBGLDVLV HOPDDKKKRGKFCFLSEYEDHKTAVQVLEFVRLAVTVEELIEKAFSQFKLERYK LNDYAFIHFDERDGAIVKEMENKQLEBENIEIVAKPDDQKREKRSQKQRAKQNM YDYVYYGPPHMPPTPRGRGRGGRGYPDYGYEDYDYGYDYHYHVRGGYEDPY YGYEDFOVGARGRGRGARGAARPRGGAAPRGAGYSGRGRGARGAARGAARGAO QQRGRGGRGVEGEPDLPQ"

BASE COUNT 1055 a 569 c 712 g 983 t  
 ORIGIN

Query Match 79.4%; Score 2741.4; DB 9; Length 3319;  
 Best Local Similarity 92.9%; Pred. No. 0;  
 Matches 3108; Conservative 0; Mismatches 106; Indels 133; Gaps 17;

QY 67 CAGGAGACAGCAGCAGCGTTCCTCCCGCAGCCGAGCCGGTGGAC-AGATTCTCGGCC 125  
 DB 10 CGGGAGACAGCAGCAGCGTTCCTCCCGCAGCCGATACCATCGGACAGATTCTCGGCC 69  
 QY 126 CAGCCCGCGGGGAGATCTCTGGAACATGCGTACAGAACATGTTAATGGAATGGTAC 185  
 DB 70 TCAGCCCAAGGGAGATCTCTGGAACATGCGTACAGAACATGTTAATGGAATGGTAC 129  
 QY 186 TGAAGACCCATGATCTCTCTGAGAGTATCATTCAGAAATTTTTCAGATTCCT 245  
 DB 130 TGAAGACCCATGATCTCTCTGAGAGTATCATTCAGAAATTTTTCAGATTCCT 189  
 QY 246 TGAATCTGTTTACACAGAAAGTTGCTGAAAACTAGATGAATTTTACGTTGAGGCT 305

DB 190 TGAATCTGTTTACACAGAAAGTTGCTGAAAACTAGATGAATTTTACGTTGAGGCT 249  
 QY 306 AGTTGACATAGATTTTGAATGAAGAAGTATCGAGCTTTAAAGATGTAATGAAGA 365  
 DB 250 AGTTGACATAGATTTTGAATGAAGAAGTATCGAGCTTTAAAGATGTAATGAAGA 309  
 QY 366 CGGCGATTTGGCAGTCTTCAACAGTTTAAAGACAGTATCTCTCATGTTGAGAACAA 425  
 DB 310 CGGTGATTTGGCAGTCTTCAACAGTTTAAAGACAGTATCTCTCATGTTGAGAACAA 369  
 QY 426 AAGTCCCTTTTATGTTGAGTCAATGAAGCTTACAGGACAGAGAAAAACAGGGACCA 485  
 DB 370 AAGTCCCTTTTATGTTGAGTCAATGAAGCTTACAGGACAGAGAAAAACAGGGACCA 429  
 QY 486 AAGTGAAGCTCTAGTAAGAGACCAATAGGCAAGATTAAGGCACTTTGAAAAAAG 545  
 DB 430 AAGTGAAGCTCTAGTAAGAGACCAATAGGCAAGATTAAGGCACTTTGAAAAAAG 489  
 QY 546 AGGCTACACACTGATGATCTACAGGTCAAGAGAGTATGAGAGACCACTTCAGATT 605  
 DB 490 AGGCTACACACTGATGATCTACAGGTCAAGAGAGTATGAGAGACCACTTCAGATT 549  
 QY 606 CGTTTATTCAGTCAAGCAGCTTCTGTGGCACTGAGATATTTGGGGAGATCCCAAG 665  
 DB 550 CGTTTATTCAGTCAAGCAGCTTCTGTGGCACTGAGATATTTGGGGAGATCCCAAG 609  
 QY 666 AGATCTGTTGAGAGAGAGCTTGTTCATTTTGAAGAAAGCTGACCTTATAGGATCT 725  
 DB 610 AGATCTGTTGAGAGAGAGCTTGTTCATTTTGAAGAAAGCTGACCTTATAGGATCT 669  
 QY 726 TCGTTTAAATGATGATCCGCTCACTGATCTCAACAGAGGTATGCTGTCATTTTGG 785  
 DB 670 TCGTTTAAATGATGATCCGCTCACTGATCTCAACAGAGGTATGCTGTCATTTTGG 729  
 QY 786 TACAAAAGAAAGCAGCAAGAGGCTGTAACTGTATTAATCATGAATTCGTTCCGG 845  
 DB 730 TACAAAAGAAAGCAGCTCAGAGAGGCTGTAACTGTATTAATCATGAATTCGTTCCGG 789  
 QY 846 GAAGCAATTTGTTGTTGATCTCACTTTCGCAATTAAGCTTTTGGGCTTCGATTC 905  
 DB 790 AAAACATTTGTTGTTGATCTCACTTTCGCAATTAAGCTTTTGGGCTTCGATTC 849  
 QY 906 TAAGATTAAGCAAGAGCAGATCTTTGAGAAATTTAGCAAGTACAGAGGCTCTCAC 965  
 DB 850 TAAGATTAAGCAAGAGCAGATCTTTGAGAAATTTAGCAAGTACAGAGGCTCTCAC 909  
 QY 966 AGATGTCATTTTATACCAACCACTGATGACAGAAAAAACAAGAGCTTTTGTCT 1025  
 DB 910 AGATGTCATTTTATACCAACCACTGATGACAGAAAAAACAAGAGCTTTTGTCT 969  
 QY 1026 TGAATTAAGATCACAAAACAGTCCCAAGCAAGAGCTAATGATGCTAAAGT 1085  
 DB 970 TGAATTAAGATCACAAAACAGTCC----- 998  
 QY 1086 CAAAGCTCGGGGAAATGTTGAACTGTTGAGTGGCTGATCTTATGAAGATCTGATCC 1145  
 DB 999 ----- 998  
 QY 1146 TGAAGTATGCAAGAGTAAAGTCTGTTTGTAGCAACTTTGCCAACAAGTAAACAG 1205  
 DB 999 -----AGTAAAGTCTGTTTGTAGCAACTTTGCCAACAAGTAAACAG 1044  
 QY 1206 AGAATTTTAAAGAAAGTCAATTAAGTATGTTGGGAAACTGGAACGATTAAGAGCTTAA 1265  
 DB 1045 AGAATTTTAAAGAAAGTCAATTAAGTATGTTGGGAAACTGGAACGATTAAGAGCTTAA 1104  
 QY 1266 AGATTAATGCTTTCATTCATTTGATGAGAGATGAGTGTCTGCAAGGCTATGGAAGAAAT 1325  
 DB 1105 AGATTAATGCTTTCATTCATTTGATGAGAGATGAGTGTCTGCAAGGCTATGGAAGAAAT 1164  
 QY 1326 GAATGCTTAAAGCTTGAAGAGAGAAAAATTAAGAAATGTTTGTCTAAGCAACAGATCA 1385  
 DB 1165 GAATGCTTAAAGCTTGAAGAGAGAAAAATTAAGAAATGTTTGTCTAAGCAACAGATCA 1224





QY	2154	TCGTGCTCAAGAGATGATTTTTCCTCCCTCTTTTAAATTAAAGATCTCAGCAGTGTATCTG	2213
Db	1854	TCGTGCTCAAGAGATGATTTTTCCTCTTTTAAATTAAAGATCTCAGCAGTGTATTTTG	1892
QY	2214	AAATATCTCTGGAATTTGTGCATTTTAAATTTTATGCAAGTATCAG---ATGCACAT	2269
Db	1893	AAATATCTCTGGAATTTGTGCATTTTAAATTTTATGCAAGTGTATTAAGATATGCCACT	1952
QY	2270	GTGGGTACCCCTAAATTTTATTTCTGCTCACCAGGTTATATCATGATGTGCTATATCTT	2339
Db	1953	GATGGTATCCCTTAAAT-TTTTATTTCTGCTCACCAGAGGTTATATCAATGATGCTATATCTT	2011
QY	2330	TTTTATGATTAATCACTTTTGAATTGNTGTAG--ATGCAGTTTCAGGTGTATATCATCAGA	2387
Db	2012	TTTTATGATGATCACTTTTGAATTGTATGTCAATATGCAAGTTTCAGGTGTATATCATCAGA	2071
QY	2388	GCTGGTATGTCAGGCAATTCACAGATAGTGATCTTTTCAGAACTCTTTT-AAAGGGTGGT	2446
Db	2072	GCTGGTATGTCAGGCAATTCAGATATAGTGATCTTTTCAGAACTCTTTTAAAGGGTGGT	2131
QY	2447	TAACTACTCAGTACGACAGAGATTTGAATCTATACCTGTCTGTATCTGTACATGAAAAATCT	2506
Db	2132	TAACTACTCAGTACGACAGAGATTTGAATCTATACCTGTCTGTATCTGTACATGAAAAATCT	2191
QY	2507	TTGTAGATPAAAAGCAAGGCTTTGTTAAATATGATATAGAGGGTATGAAATTTTAAATATCCAA	2566
Db	2132	TTGTAGATPAAAAGCAAGGCTTTGTTAAATATGATATAGAGGGTATGAAATTTTAAATATCCAA	2251
QY	2567	TGTAACTTCTTACGTGCTTGTATGATTTCAAGAGCTTGTAAAGACTTCCATCATGACCATAT	2626
Db	2252	TGTAACTTCTTACGTGCTTGTATGATTTCAAGAGCTTGTAAAGACTTCCATCATGACCATAT	2311
QY	2627	AACGAGCTTGCTTTGTGCTATTTTGTGGCTGAATAAGCAGCCTTGCTCTTCAGATAT	2686
Db	2312	AACGAGCTTGCTTTGTGCTATTTTGTGGCTGAATAAGCAGCCTTGCTCTTCAGATAT	2371
QY	2687	TGTAGTATTTTGANTGTATATATGTTTAGAAGATGTATCTTTGTAAAGACTCAGATGT	2746
Db	2372	TGTAGTATTTTGANTGTATATATGTTTAGAAGATGTATCTTTGTAAAGACTCAGATGT	2431
QY	2747	TCAAAAAAAGTGCATCCGCACTTGTAATATCTGACAGTGCCTCTTATPAAAAAGTGC	2806
Db	2432	TC-----AAAAAGTGCATCCGCACTTGTAATATCTGACAGTGCCTCTTATPAAAAAGTGC	2488
QY	2807	AGACTAAACTGACATTTGTACAGCAAAAGCTGACATTTGATATTTTGAAGTTTTTCA	2866
Db	2489	AGACTAAACTGACATTTGTATGTATGTATAGACGAAAGCTGACATTTGATATTTTGAAGTTTTTCA	2548
QY	2867	TAAATCAATGAATAATATATATATATGAGCTGTAGTTTACCTTTTAGTAAAGATAGTTTCA	2926
Db	2549	TAAATCAATGAATAATATATATATATGAGCTGTAGTTTACCTTTTAGTAAAGATAGTTTCA	2608
QY	2927	TTATGCACTTTGTTATATGCTGATCATCTATPAAAAATGTAAATCAGGTTCCATTTCTATAG	2966
Db	2609	TTATGCACTTTCTTCCTGCTGATCATCTATPAAAAATGTAAATCAGGTTCCATTTCTATAG	2668
QY	2967	CAGGTCAATGATTAACCTTGTAGATATAGATACAATCATTTGTGTATGTTTTTAAATTTCTA	3046
Db	2669	CAGGTCAATGATTAACCTTGTAGATATAGATACAATCATTTGTGTATGTTTTTAAATTTCTA	2728
QY	3047	AAGCACTTGATGACATATGATGTTTCAGTGTGTAGAGCAATCTCTCTATGAAATCCCTCAA	3106
Db	2729	AAGCACTTGATGACATATGATGTTTCAGTGTGTGTAGAGCAATCTCTCTATGAAATCCCTCAA	2787
QY	3107	AAAAATTTTGTGCAAGTCTCAAGTTGATATAGCTTAAAGTCAAAAGT-AAAATTTATAGTTT	3165
Db	2788	AAAAATTTTGTGCAAGTCTCAAGTTGATATAGCTTAAAGTAAAGTGAATAATATATAGTTT	2847
QY	3166	AAGTAGACATTGTGTATAAGAAAACACCTT-CCGCCCTTCCCAAGAGATACGTGACGATGCT	3224
Db	2848	CATTAGACATTGTGTATAAGAAAATCCCTCCCTCCCTCCCAAGAGATATACGTGACGATGAT	2907
QY	3225	ATTCACACCCACATAGAGCACACAGATATGAGATCAGAC-TAGACCTGACTAAGGTTTAT	3283

Db	2908	ATCACATACCCCAATGAGCAACCAAGTGAAGATCAGAGCTTACTTAATTGAAGTTTAT	2967
Qy	3284	ACACACACAGTTCACCAATGAATGCAATTTTAAACAAGAAATTAGACATGATATGTTCA	3343
Db	2968	ACACACACAGTTCACCAATGAATGCAATTTTAAACAAGAAATTAGACATGATATGTTCA	3027
Qy	3344	AAATGCTCATGGCAACAATCATTTTGCATTCCTGCAATATAATTTGTTTATAC	3398
Db	3028	AAATGCTCATGGCAACAATCATTTTGCATTCCTGCAATATAATTTGTTTATAC	3082
RESULT 4			
LOCUS	AF155568	2221 bp	mRNA linear PRI 10-JUN-1999
DEFINITION	Homo sapiens NSAP1 protein (NSAP1)	complete cds.	
ACCESSION	AF155568		
VERSION	AF155568.1	GI:5031511	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 2221)		
TITLE	Harris,C.E., Boden,R.A. and Astell,C.R.		
JOURNAL	A novel heterogeneous nuclear ribonucleoprotein-like protein		
MEDLINE	interacts with NS1 of the minute virus of mice		
PIRME	J. Virol. 73 (1), 72-80 (1999)		
REFERENCE	9847309		
AUTHORS	2 (bases 1 to 2221)		
TITLE	Harris,C.E., Boden,R.A. and Astell,C.R.		
JOURNAL	Submitted (01-JUN-1999) Biochemistry and Molecular Biology, UBC,		
FEATURES	2146 Health Sciences Mall, Vancouver, BC V6T 1Z3, Canada		
SOURCE	Location/Qualifiers		
gene	1..2221		
CDs	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/note="cell line"		
	1..2221		
	/gene="NSAP1"		
	205..1893		
	/gene="NSAP1"		
	/note="similar to Homo sapiens hn RNP R; probable RNA		
	binding protein; interacts with NS1 of the minute virus of		
	mice"		
	/codon_start=1		
	/product="NSAP1 protein"		
	/protein_id="A038198.1"		
	/db_xref="GI:5031512"		
	/translation="MATEHNKNGTSEPMDTTSVAIVHSERFQTLDAELPQKTAELKD		
	ELIVAGLVASDDDERALTEKFNEDGALAVIQFSDLSHVOKSAFLCGMKTLY		
	ROEKQSTKAVSDSKGPDKEAKIKALIERGYTLDDVTGQRKGSGPPDSVYSGQPSV		
	GTEIYFKELPRDLFEDLELPIFEKAGPIWDLRLMDPLFGLNKGVAFTFCREAQOE		
	AVGLYNNHEIRSGKHIGVCSIVANNRLFVGSIPKSTKQQLIEEPQVTEGELTDVILY		
	HODDDKKKNGRGPELEVEDHKTAQAQRRLMSKRYVMGVGVEMADPIEDDEPYM		
	AKTKVLFYRRLATVTYBELIERKASQFGKILRVKLDYAFHFEDRDGAVKMEEN		
	KGKLEGENLEIVFAKPPDQRRKERAQQAANKQMDVYYVGPPIIMPETRRGSG		
	RGGYGAPPDYGVEDYDYGVYHVRGGYEDPYGVYGVDFGVGAAGGGGGRGARGAP		
	SRRGGAAPPRGARGYSGRGGPGSARGVARGAGAAQQRGGQKGVAGPDLLQ"		
BASE COUNT	699 a 402 c 561 g 559 t		
ORIGIN			
Query Match	56.6%;	Score 1953.8;	DB 9; Length 2221;
Best Local Similarity	95.6%;	Pred. No. 0;	
Matches 2063; Conservative	0;	Mismatches 87;	Indels 7; Gaps 5;
Qy	2	CGGGTTGCTGTCGAGCGGCGTGAAGCTTCGCGCGCATTTTAAACAAGCTCACAATCGGCGC	61
Db	52	CGGGTTGCTGTCGAGCGGCGTGAAGCTTCGCGCGCATTTTAAACAAGCTCACAATCGGCGC	111
Qy	62	GGAACAAGGAGCAAGCAAGCGTTTCCCGCGCAAGCCGAGTGGAGC-GGAATTCCT	120









QY 1201 ACAGAGAAATTTAGAAAGTCAATTTAGTCAATTTGGAAACTGGAACGATTAAG 1260  
 Db 1250 ACAGAGAGATTTAGAAAGGATTTAGTCAATTTGGAAACTGGAACGATTAAG 1309  
 QY 1261 CTAAAGATTTATGCTTTTCAATTTGATGAGAGATGCTGCTCAAGGCTATGAA 1320  
 Db 1310 TTTAAAGATTTATGCTTTTCAATTTGATGAGAGATGCTGCTCAAGGCTATGAA 1369  
 QY 1321 GAAATGATGTTAAAGCTTGGAGGAGAAATTTGAAATTTGTTTGTGTAAGCCCA 1380  
 Db 1370 GAAATGATGTTAAAGCTTGGAGGAGAAATTTGAAATTTGTTTGTGTAAGCCCA 1429  
 QY 1381 GATCAGAGAGAGAAAGAAAGAAAGCTCAGAGGCAAGCAAGAAATCAATGATAT 1440  
 Db 1430 GATCAGAGAGAGAAAGAAAGAAAGCTCAGAGGCAAGCAAGAAATCAATGATAT 1489  
 QY 1441 GATTACTATTTATTTGCTTCCATCATATGCTCTCCCAACAGAGGTGAGGGCTGGA 1500  
 Db 1490 GATTACTATTTATTTGCTTCCATCATATGCTCTCCCAACAGAGGTGAGGGCTGGA 1549  
 QY 1501 GGTAGAGGTGCTATGATATCTCCAGATTTATGATATGATATGATAT 1560  
 Db 1550 GGTAGAGGTGCTATGATATCTCCAGATTTATGATATGATATGATAT 1609  
 QY 1561 TATGTTATGATTAACATTAATCTATGCTGATATGATATGATATGATAT 1620  
 Db 1610 TATGTTATGATTAACATTAATCTATGCTGATATGATATGATATGATAT 1669  
 QY 1621 GATTTTCAATTTGAGCTTAAGAGAGAGGGGTGATAGAGAGAGGGGTGCTCTCATCC 1680  
 Db 1670 GATTTTCAATTTGAGCTTAAGAGAGAGGGGTGATAGAGAGAGGGGTGCTCTCATCC 1729  
 QY 1681 AGAGGTGCTGAGGGGTGCTCTCCAGATTTATGATATGATATGATATGATAT 1740  
 Db 1730 AGAGGTGCTGAGGGGTGCTCTCCAGATTTATGATATGATATGATATGATAT 1789  
 QY 1741 GATCAGAGAGAGAGAGAGAGGGGTGATAGAGAGAGGGGTGCTCTCATCC 1800  
 Db 1790 GATCAGAGAGAGAGAGAGAGGGGTGATAGAGAGAGGGGTGCTCTCATCC 1849  
 QY 1801 GGAAGAGGGGTGAGAGGGGTGCTCTGATCATGAAAGACTGATGATGATGAG 1860  
 Db 1815 GGAAGAGGGGTGAGAGGGGTGCTCTGATCATGAAAGACTGATGATGAG 1873  
 QY 1861 ATTACACAGAGAGAGAGAGAGAGGGGTGATAGAGAGAGGGGTGCTCTCATCC 1920  
 Db 1874 ATTACACAGAGAGAGAGAGAGAGGGGTGATAGAGAGAGGGGTGCTCTCATCC 1932  
 QY 1921 CGGATGATAGAGAGAGAGAGAGAGGGGTGATAGAGAGAGGGGTGCTCTCATCC 1980  
 Db 1933 CGGATGATAGAGAGAGAGAGAGAGGGGTGATAGAGAGAGGGGTGCTCTCATCC 1992  
 QY 1981 TAGGAGAGAGAGAGAGAGAGGGGTGATAGAGAGAGGGGTGCTCTCATCC 2040  
 Db 1993 TAGGAGAGAGAGAGAGAGAGGGGTGATAGAGAGAGGGGTGCTCTCATCC 2052  
 QY 2041 AGTCAAGTATTAAGTAAAGTACT 2067  
 Db 2053 AGTCAAGTATTAAGTAAAGTACT 2079

RESULT 7  
 LOCUS HSJ775C13 103819 bp DNA linear PRI 10-MAR-2001  
 DEFINITION Human DNA sequence from clone RP4-775C13 on chromosome 20p12.1-13.  
 Contains an hnRNP R (heterogenous nuclear ribonucleoprotein R, RRM rna  
 binding protein GRP-RBP) pseudogene, ESTs, STS and GSS, complete  
 sequence.  
 ACCESSION AL109618 GI:5668665  
 VERSION AL109618.1  
 KEYWORDS HTG; GRP-RBP; hnRNP; RRM.  
 SOURCE human.  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 103819)  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Submitted (02-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 requests: clonerequest@sanger.ac.uk  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information on the WORMPEP  
 database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)  
 This sequence was generated from part of bacterial clone contigs of human  
 chromosome 20, constructed by the Sanger Centre Chromosome 20  
 Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr20>  
 RP4-775C13 is from the library RPCI-4 constructed by the group of  
 Pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pCYPAC2

#### FEATURES

source  
 1..103819  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="p12.1-13"  
 /clone="RP4-775C13"  
 /clone\_1fb="RPCI-4"  
 1..236  
 /note="AluY repeat: matches 1..236 of consensus"  
 293..562  
 /note="AluMc repeat: matches 1844..1624 of consensus"  
 819..1471  
 /note="AluMc repeat: matches 969..1634 of consensus"  
 2177..2588  
 /note="AluMc repeat: matches 14..431 of consensus"  
 2593..2652  
 /note="AluMc repeat: matches 474..533 of consensus"  
 2701..3176  
 /note="238 copies 2 mer tt 54% conserved"  
 2814..2978  
 /note="15 copies 11 mer 66% conserved"  
 2996..3196  
 /note="67 copies 3 mer tcc 56% conserved"  
 3031..3184  
 /note="14 copies 11 mer 63% conserved"  
 3296..4082  
 /note="AluMc repeat: matches 3350..4128 of consensus"  
 4197..4315  
 /note="AluMc repeat: matches 3021..3136 of consensus"  
 4316..4738  
 /note="MSTR repeat: matches 1..426 of consensus"  
 4739..5118  
 /note="AluMc repeat: matches 2626..3021 of consensus"  
 5146..5288  
 /note="match: GSS; Em:B90786"  
 5345..6619  
 /note="AluA6 repeat: matches 3863..5135 of consensus"



QY 362 AAGACGCGCATTTGGCAGTCTTCAACAGTTTAAAGACAGTGTCTCTCATGTTGAGA 421  
 DB 7559 AAGACGCGCATTTGGCAGTCTTCAACAG--TTAAAGACAGTGTCTCTCTGTTGAGA 7617  
 QY 422 ACAAAAGTGCCTTTTATGTGAGTCAATGAAGCTTAACAGGAGAGAAACAGGGGA 481  
 DB 7618 ACAAAAGTGCCTTTTATGTGAGTCAATGAAGCTTAAAGGAGAGAAACAGGGGA 7677  
 QY 482 CCAAGTAGCAGCTCTAGTAAAGAGACCAATGAGGCAAGATTAAAGGCACTTTGAAA 541  
 DB 7678 CCAAGTAGCAGTCTCTAGTAAAGAGACCAATGAGGCAAGATTAAAGGCACTTTGAAA 7737  
 QY 542 GAACAGGCTACACATTTAGTGAAGTCAAGTCAAGAGAGATAGAGAGCACTTCAG 601  
 DB 7738 GAACAGGCTACACATTTAGTGAAGTCAAGTCAAGAGAGATAGAGAGCACTTCAG 7797  
 QY 602 ATTCCGTTTATTTAGGTCAGAGCTTCTGTTGGCACTGAGATATTTGTGGGAAAGATCC 661  
 DB 7798 ATTCCGTTTATTTAGGTCAGAGCTTCTGTTGGCACTGAGATATTTGTGGGAAAGATCC 7857  
 QY 662 CCAAGATCTGTTGAGGATGAGCTTTCATTAATTTGAGAAAGCTGAGCTATATGAG 721  
 DB 7858 CCAAGATCTATTTGAGGATGAGCTTTCATTAATTTGAGAAAGCTGAGCTATATGAG 7917  
 QY 722 ATCTTGTTTAAATGATGATCCGCTCACTGATCTCAACAGAGTTATGCTTTGCACTT 781  
 DB 7918 ATCTTGTTTAAATGATGATCCGCTCACTGATCTCAATAGGTTATGCTTTGCACTT 7977  
 QY 782 TTTGTACAAAAGAGCAGCACAAGAGGCTGTTAACTGTATTAATCATGAATTCGTT 841  
 DB 7978 TTTGTACAAAAGAGCAGCACAAGAGGCTGTTAACTGTATTAATCATGAATTCGTT 8035  
 QY 842 CGGGGAAGCAGATGAGTGTGATCTCACTGATCTCAACATTAAGCTTTTGTGGCTGCA 901  
 DB 8036 CGGGGAAGCAGATGAGTGTGATCTCACTGATCTCAACATTAAGCTTTTGTGGCTGCA 8095  
 QY 902 TTCTTAAGAG-TAAACCAAGAGCAGATCTTGAAGGATTTAGC-AAATGACAGAGG 959  
 DB 8096 TTCTTAAGAGTTAAACCAAGAGCAGATCTTGAAGGATTTAGGCAAAATTAACAGAGG 8155  
 QY 960 TCTCAAGATGTATTTTATACCAACCACTGATGACAGAGAAACAGAGGCTTTTG 1019  
 DB 8156 TCTCAAGATGTATTTTATACCAACCACTGATGACAGAGAAACAGAGGCTTTTG 8213  
 QY 1020 CTTTCTTGAATGAGATGACAAACAGCTGCCAGGCAAGCTATGAGTGTG 1079  
 DB 8214 CTTTCTTGAATGAGATGACAAACAGCTGCCAGGCAAGCTATGAGTGTG 8273  
 QY 1080 TAAAGTCAAAGTCTGGGGAATGTTGGAATGTTGAGTGGGCTGATCTATGGAAGATCC 1139  
 DB 8274 TAAAGTCAAAGTCTGGGGAATGTTGGAATGTTGAGTGGGCTGATCTATGGAAGATCC 8333  
 QY 1140 TGAATCTGAAGTTATGAGCAAGGTAAGTGTGTTTGAACGAACTTTGCCAACAGCT 1199  
 DB 8334 TGAATCTGAAGTTATGAGCAAGGTAAGTGTGTTTGAACGAACTTTGCCAACAGCT 8393  
 QY 1200 AACAGAAAGAAATTTTGAAGAGTCACTTACTGAGTTGGGAACTGGAAGATGAAGAA 1259  
 DB 8394 AACAGAAAGAAATTTTGAAGAGTCACTTACTGAGTTGGGAACTGGAAGATGAAGAA 8453  
 QY 1260 GCTTAAAGATTAATGCTTCACTTATTTGATGAGAGATGAGTGTCTCAAGGCTATGGA 1319  
 DB 8454 GCTTAAAGATTAATGCTTCACTTATTTGATGAGAGATGAGTGTCTCAAGGCTATGGA 8513  
 QY 1320 AGAATGATGTGAAGCTTGAAGAGGAAATATTAATGAAATGTTTGTCTAAGCCACC 1379  
 DB 8514 AGAATGATGTGAAGCTTGAAGAGGAAATATTAATGAAATGTTTGTCTAAGCCACC 8573  
 QY 1380 AGATCAAGAGAGAAAGAAAGAAAGCTCAGAGCAGAGCAAGATCAATATGATGA 1439  
 DB 8574 AGATCAAGAGAGAAAGAAAGAAAGCTCAGAGCAGAGCAAGATCAATATGATGA 8631  
 QY 1440 TGAATCACTATTAATGATGATCACTCATATGCTCCCAACAGAGGTGAGAGGCTG 1499

DB 8632 CGATTACTATTAATGATGATCACTTATATGCCCTCCAAAGAGATGAGAGGAGG 8691  
 QY 1500 AGGTAGAGGTGCTATGATGATATCTCCAGATTTATTAAGATGAAGATTAATGATTA 1559  
 DB 8692 AGGTAGAGGTGCTATGATGATATCTCCAGATTTATTAAGATGAAGATTAATGATTA 8751  
 QY 1560 TTAATGATTAATGATTAATCACTATGATGATGATGATGATGATGATGATGATGATGAT 1619  
 DB 8752 TTAATGATTAATGATTAATCACTATGATGATGATGATGATGATGATGATGATGATGAT 8811  
 QY 1620 AGATTTTCAAGTTGAGCTAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1679  
 DB 8812 AGATTTTCAAGTTGAGCTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8870  
 QY 1680 CAGAGGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1739  
 DB 8871 CAGAGGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8930  
 QY 1740 TGAATGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1799  
 DB 8931 TGAATGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8990  
 QY 1800 ----GGGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1856  
 DB 8991 GAGG 9049  
 QY 1857 TGGGATTAACCAAGAGCTTGCAGTGAAGTAAATGATGAAGAAATCAACAGCACTTAAAT 1916  
 DB 9050 TGGGATTAACCAAGAGCTTGCAGTGAAGTAAATGATGAAGAAATCAACAGCACTTAAAT 1918  
 QY 1917 ATCTGAGATTAATGAGAGCATATTTGAGAGAAACCTCTATGAAGATGATGAAT 1976  
 DB 9109 ATCTGAGATTAATGAGAGCATATTTGAGAGAAACCTCTATGAAGATGATGAAT 1968  
 QY 1977 CAAATAGGAGCATTAATTAATCTTGAAGCTTGTATGAATTTCTTTTAAACAAATTTCT 2036  
 DB 9169 CAAATAGGAGCATTAATTAATCTTGAAGCTTGTATGAATTTCTTTTAAACAAATTTCT 9228  
 QY 2037 CTGAGTGAAGTAAATTAATTAAGCTTACTATTTTCCAAATGTTCCAA-AAAAAT 2095  
 DB 9229 CTGAGTGAAGTAAATTAATTAAGCTTACTATTTTCCAAATGTTTCCAAAGAAAT 9288  
 QY 2096 CTTTCACTTCTTACATGATCTTATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 2151  
 DB 9289 CTTTCACTTCTTACATGATCTTATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 9344

RESULT 8  
 AL390737 212659 bp DNA linear HTG 21-MAR-2001  
 LOCUS  
 DEFINITION Homo sapiens chromosome 13 clone RP11-190P17, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\*, 16 unordered pieces.  
 ACCESSION AL390737  
 VERSION AL390737.9 GI:13398796  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 AUTHORS Burton, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 requests: clonequery@sanger.ac.uk  
 On Mar 21, 2001 this sequence version replaced gi:13398612.

COMMENT  
 ----- Genome Center  
 Center: Sanger Centre  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquery@sanger.ac.uk  
 ----- Project Information

		Query Match	Best Local Similarity	Matches 2014; Conserveative	52.2%; Pred. No. 0; Mismatches 125; Indels 17; Gaps 12;	Score 1802; DB 2; Length 212659;
	misc_feature	/note="assembly_fragment:00015"	171360..173612			
	misc_feature	/note="assembly_fragment:00477"	173713..187445			
	misc_feature	/note="assembly_fragment:00780"	187546..191084			
	misc_feature	/note="assembly_fragment:01634"	191185..194709			
	misc_feature	/note="assembly_fragment:01673"	194810..197824			
	misc_feature	/note="assembly_fragment:02018"	197925..200623			
	misc_feature	/note="assembly_fragment:02519"	200724..203433			
	misc_feature	/note="assembly_fragment:02839"	203534..206523			
	misc_feature	/note="assembly_fragment:02989"	206624..208875			
	misc_feature	/note="assembly_fragment:03934"	208976..212659			
	misc_feature	/note="assembly_fragment:04063"				
	BASE COUNT	61246 a 41295 c 42666 g 65947 t 1505 others				
	ORIGIN					
Qy	3	GGGTTTGCTGCAGCGGGGTGAGCTTCGGCCGCACTTTTACAGAGCTCACTCGCCG	62			
Db	82955	GGGTTTGCTGCAGCGGGGTGAGCTTCGGCCGCACTTTTACAGAGCTCACTCGCCG	83014			
Qy	63	GACACAGGAGCAGCGAGCAGCGCTCCCGCAGCCGACCCGGTCGAC-GGATTCCTC	121			
Db	83015	GACACAGGAGCAGCGAGCAGCGCTTCCTTCGACCTGTCATCGCAGAGATTCCTC	83074			
Qy	122	GGCCCCAGCCCCGGGGGAGATCTCTGAAACATGGCTACGAACTATATGAAATG	181			
Db	83075	CCCCCACTCCACGGGGAGATCTTGGAAACATGGCTACGAACTATATGAAATG	83134			
Qy	182	GTACTGAAGAGCCCATGATCTACTCTCAGAGATTATCATTCAGAAAATTTGACAT	241			
Db	83135	ATACGAAAGCCCATGATCTACTCTCAGAGATTATCATTCAGAAAATTTGACAT	83194			
Qy	242	TGCTGATGCTGCTTTCACAGAAAGTTGCTGAAAACCTAGAGAAATTTACGTTGAG	301			
Db	83195	TGCTGATGCTGCTTTCACAGAAAGTTGCTGAAAACCTAGAGAAATTTACGTTGAG	83254			
Qy	302	GGCTAGTTCACATATGATTTAATGATAAAGACTATGAAAGCTTTAAAGATTCAATG	361			
Db	83255	GGCTAGTTCACATATGATTTAATGATAAAGACTATGAAAGCTTTAAAGATTCAATG	83314			
Qy	362	AAGAGCGGCGCATTTGGCAGTGTCTTCAAGCTTTAAAGACAGATCTCTCATGTTGCA	421			
Db	83315	AAGAGCGGCGCATTTGGCAGTGTCTTCAAGCTTTAAAGACAGATCTCTCTGTTTCA	83373			
Qy	422	ACAAAAGTCCCTTTTATGAGATCTAGAAAGCTTACAGCGCAGAGAAAACAGGGGA	481			
Db	83374	ACAAAAGTCCCTTTTATGAGATCTAGAAAGCTTATAGCAGAGAAAACAGGGGA	83433			
Qy	482	CCAAAAGTGCAGATCTTATGATTAAGAACAGATGAGGCAAAAATTTAAGCCTTTGAAA	541			
Db	83434	CCAAAAGTGCAGATCTTATGATTAAGAACAGATGAGGCAAAAATTTAAGCCTTTGAAA	83493			
Qy	542	GAAAGGCTAACACTTGAATGTGACTACAGTCAAGAAATATGAGAGACCACTCCAG	601			
Db	83494	GAAAGGCTAACACTTGAATGTGACTACAGTCAAGAAATATGAGAGACCACTCCAG	83553			
Qy	602	ATTCGTTTATTCAGGTGAGAGCCTTCTGTTGGCACTGAGATATTTGTGGGAAAGATCC	661			
Db	83554	ATTCGTTTATTCAGGTGAGAGCCTTCTGTTGGCACTGAGATATTTGTGGGAAAGATCC	83613			
Qy	662	CCAGAGATCTGTTGAGGATGAGCTTGTTCATTTATTTGAGAAAAGCTGACCTATATGG	721			



Db 83614 CAGAGATCATTTGAGAGTGAAGTCTGTCATTTATTTGAGAAAGCTGCTATATGAG 83673  
 Qy 722 ATCTTGCTTTAATGATGATCCGCTCACTGCTCTCAACAGAGTTATGCTTTGCTACTT 761  
 Db 83674 ATCTTGCTTTAATGATGATCCGCTCACTGCTCTCAACAGAGTTATGCTTTGCTACTT 83733  
 Qy 782 TTTGTCAAAAG 841  
 Db 83734 TTTGTCAAAAG 83791  
 Qy 842 CCGGGAAGACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 901  
 Db 83792 CTGGAAGAACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 83851  
 Qy 902 TTTCTTAAGAG-TAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 959  
 Db 83852 TTTCTTAAGAG-TAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 83911  
 Qy 960 TCTCAGAGATGCTATTTATACCAACCACTGATGACCAAGAGAGAGAGAGAGAGAG 1019  
 Db 83912 TCTTACAGAGATGCTATTTATACCAACCACTGATGACCAAGAGAGAGAGAGAG 83969  
 Qy 1020 CTTTCTTAAGATGAGATCAACCAACAGCTGCTCAGGCAAGAGAGAGAGAGAGAG 1079  
 Db 83970 CTTTCTTAAGATGAGATCAACCAACAGCTGCTCAGGCAAGAGAGAGAGAGAG 84029  
 Qy 1080 TAAAGTCAAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1139  
 Db 84030 TAAAGTCAAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 84089  
 Qy 1140 TGAATCTGAAGTATGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1199  
 Db 84090 TGAATCTGAAGTATGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 84149  
 Qy 1200 AACAG 1259  
 Db 84150 AACAG 84209  
 Qy 1260 GCTTAAAGATGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1319  
 Db 84210 GCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 84269  
 Qy 1320 AGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1379  
 Db 84270 AGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 84329  
 Qy 1380 AGATCAG 1439  
 Db 84330 AGATCAG 84387  
 Qy 1440 TGAATCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1499  
 Db 84388 TGAATCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 84447  
 Qy 1500 AGGTAG 1559  
 Db 84448 AGGTAG 84507  
 Qy 1560 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1619  
 Db 84508 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 84567  
 Qy 1620 AATTTTCAAGTTGAG 1679  
 Db 84568 AATTTTCAAGTTGAG 84626  
 Qy 1680 CAGAGAGTCTGAG 1739  
 Db 84627 CAGAGAGTCTGAG 84686  
 Qy 1740 TGAATCAG 1799

Db 84687 TGAATCAG 84746  
 Qy 1800 ---GGGAAAAG 1856  
 Db 84747 GAG 84805  
 Qy 1857 TGGGATTTACAG 1916  
 Db 84806 TGGGATTTACAG 84864  
 Qy 1917 ATCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1976  
 Db 84865 ATCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 84924  
 Qy 1977 CAATTCAG 2036  
 Db 84925 CAATTCAG 84984  
 Qy 2037 CTGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2095  
 Db 84985 CTGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 85044  
 Qy 2096 CTTTCAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2151  
 Db 85045 CTTTCAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 85100  
 RESULT 9  
 AF037448 2932 bp mRNA linear PRI 08-APR-1998  
 LOCUS AF037448  
 DEFINITION Homo sapiens RRM RNA binding protein GRY-1BP (GRY-RBP) mRNA,  
 complete cds.  
 ACCESSION AF037448  
 VERSION AF037448.1 GI:3037012  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 2932)  
 AUTHORS Du G., Yuan J., Pan M., Yao H., Chen J. and Qiang B.  
 TITLE Cloning of a RRM RNA binding protein with glycine, arginine and  
 tyrosine-rich C-terminus  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 2932)  
 AUTHORS Du G., Yuan J., Pan M., Yao H., Chen J. and Qiang B.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-DEC-1997) Department of Biochemistry and Molecular  
 Biology, Institute of Basic Medical Sciences, 5 Dong Dan San Tiao,  
 Beijing 100005, P.R. China  
 FEATURES  
 source  
 1..2932  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 1..2932  
 /gene="GRY-RBP"  
 526..2397  
 /gene="GRY-RBP"  
 /note="RRM RNA binding protein; RNA recognition motif  
 binding protein; Glycine, arginine and tyrosine-rich at  
 C-terminus"  
 /codon\_start=1  
 /product="Gry-1bp"  
 /protein\_id="AAC12926.1"  
 /db\_xref="GI:3037013"  
 /translation="MATEHVNGNTEBPMDTSAVHSNPTLLDGLPKVAKLID  
 EIVAGLVASHDLDEPALKERNEGALAVLQFPDSDLSHKNSAPLCTGKTY  
 ROKRQCTKVAADSKDEPKIKALIRGTGLDVLVGGRTGSPDVSYSQOQSV  
 GTEIFVGIKIPDLFEDLVLPFERAGIPWDIRLMDPLVGLNRGAVAFCTKXAOE  
 AVKLYNHHEISGHIIGVCSVANNRIFVGSIPSKTKEQILLESFSTVTEGLDVLVY  
 HOPDKKKNRSFCELEYEDHKTAAQARRRLSGKVGKVGAVTEMADPIEDPPEVY  
 AKVAVLFRNLANTVTEILLEKASQSGKLERVKLQDVAFIHFDERDGAVKAMEEN  
 GKLEGENIEIVPAKPPDQKREKKAQQAANKQMDYDYYTGPPHPPPTRGGRGG

BASE COUNT	868 a	520 c	823 g	721 t	ORIGIN
					RGCGGYPDPDYGYEDYYDYVYDYNHNRGGTJEDDPYTGIEDYQVAGARGGRGARGAAP
					SRGGCAAPRRAGAGSQRGGPSSAGATGTCGAGGAQOQQRGRVRCARGGRGAGNVGAKRK
					ADGNNQPDPSKRRQTNNQNWGSQPTAQQDLOGSDHSNNTGYSENQFQDTDFGQMK"

Query Match	48.5%	Score 1674;	DB 9;	Length 2932;
Beat Local Similarity	95.3%;	Pred. No. 0;		
Matches 1736; Conservative	0;	Mismatches 85;	Indels 1;	Gaps 14;

OY	2	CGGGTTTCTCCAGCGCGCGGACCTTCGCGCGCAATTTTAAACAAGCTCCAGCTCGCGCC	61
Db	373	CGGGTTTCTCCAGCGCGCGGACCTTCGCGCGCAATTTTAAACAAGCTCCAGCTCGCGCC	432
OY	62	GGAACAAGGAGCAGCGAGCAACGCGTTCGCCGCAAGCCGCGTCCGAC-GGATTCCT	120
Db	433	GGAACAAGGAGCAGCGAGCAACGCGTTCGCCGCAAGCCGCGTCCGAC-GGATTCCT	492
OY	121	CGGCCCCGAGCCCCCGGGGAGATTCCTGGAAAACATGGCTACAGAACATGTTAATGGAAAT	180
Db	493	CGGCTCAGCCCAACGGGGAGATTCCTGGAAAACATGGCTACAGAACATGTTAATGGAAAT	552
OY	181	GGTACTGAAGAGCCCATGGATATCACTACTCAGCAGTTATCCATTACAGAAATTTTACAGCA	240
Db	553	GGTACTGAAGAGCCCATGGATATCACTACTCAGCAGTTATCCATTACAGAAATTTTACAGCA	612
OY	241	TTTGCTTGAATGCTGGTTTACCAACAGAAATGCTGAAAACCTAGATGAATTTTACCTTCA	300
Db	613	TTTGCTTGAATGCTGGTTTACCAACAGAAATGCTGAAAACCTAGATGAATTTTACCTTCA	672
OY	301	GGGCTTAGTTGCACTAATGATTTTATGATGAAGAAGCTATCGAAGCTTTAAAGATTTCAAT	360
Db	673	GGGCTTAGTTGCACTAATGATTTTATGATGAAGAAGCTATGGAAGCTTTAAAGATTTCAAT	732
OY	361	GAAAGCGGCGCATGGGAGAGGCTTCAACAGTTTAAAGAACAGATATCTCTCATATGTAC	420
Db	733	GAAAGCGGCGCATGGGAGAGGCTTCTTCAACAGTTTAAAGAACAGATATCTCTCATATGTAC	792
OY	421	AACAAAGATGCTTTTATATGTGAGTCAATGACATTAACGACAGAGAGAGAAAACAGGGG	480
Db	793	AACAAAGATGCTTTTATATGTGAGTCAATGACATTAACGACAGAGAGAGAAAACAGGGG	852
OY	481	ACCAAAGTAGACACTTAATGAAAGAACAGATGAGCAAGATTAAGCACTTTTGAA	540
Db	853	ACCAAAGTAGACATTTAGTAAAGAACAGATGAGCAAAATTAAGCACTTTTGAA	912
OY	541	AGAACAGGCTACACACTTGAATGAGCTACAGAGCTCAGAGGAAGATAGAGAACCACTTCA	600
Db	913	AGAACAGGCTACACACTTGAATGAGCTCAGAGCTCAGAGGAAGATAGAGAACCACTTCA	972
OY	601	GATTCCTGTTTATTCAGAGTACAGACGCTTCTGTGGCACTGAGATATTTGTGGGAAGATC	660
Db	973	GATTCCTGTTTATTCAGAGTACAGACGCTTCTGTGGCACTGAGATATTTGTGGGAAGATC	1032
OY	661	CCCAAGATCTGTTTGAAGATGAGCTTGTCCATTTATGAGAAAGCTGACCTTATATGG	720
Db	1033	CCCAAGATCTAATTTGAGATGAGACTTGTCCATTTATGAGAAAGCTGACCTTATATGG	1092
OY	721	GATCTGCTTAAATGATGAGATCCGCTCACAGGCTCAACAAGAGTTAATGCGTTTCTACT	780
Db	1093	GATCTGCTCTAATGATGAGATCCACTACAGGCTCAATAAGAGTTAATGCGTTTCTACT	1152
OY	781	TTTTTGTACAAAAGAGACAGACAAGAGGCTGTAAACTGTATATATATCATGAAATTCGT	840
Db	1153	TTTTTGTACAAAAGAGAGCGTCAAGAGGCTGTAAACTGTATATATATCATGAAATTCGT	1212
OY	841	TCCGGGAAGCATTGGTGTCTGATTCAGTTGCCAACATAGGCTTTTGTGGGCTCG	900
Db	1213	TCTGGAAAACATATTTGGTGTCTGATTCAGTTGCCAACATATAGGCTTTTGTGGGCTCT	1272
OY	901	ATTCCTTAGAGATTAACCAAGAGAGACATTCCTGAAGAAATTTAGCAAAATGACAAAGGT	960
Db	1273	ATTCCTTAGAGATTAACCAAGAGACATTCCTGAAGAAATTTAGCAAAATGACAAAGGT	1332

QY	961	CTCAGAGATGTCATTTTATACCAACCACTGATGCAAGAAAAAACAAGAGGCTTTTGC	1020
Db	1333	CTTACAGACGTCATTTTATACCAACCACTGATGCAAGAAAAAACAAGAGCTTTTGC	1392
QY	1021	TTTCTTGAATATGAAGATCAAAAACAGCTGCGCCAGCAAGCGTAGCTATAGAGTGT	1080
Db	1393	TTTCTTGAATATGAAGATCAAAAACAGCTGCGCCAGCAAGCGTAGCTATAGAGTGT	1452
QY	1081	AAAGCAAAAGCTCGGGGAAATGTTGGAACGTGTAGTGGGCTGATGCTATAGAGATCCT	1140
Db	1453	AAAGTCAAAGCTCGGGGAAATGTTGGAACGTGTAGTGGGCTGATGCTATAGAGATCCT	1512
QY	1141	GATCCTGAAGTTATGCGCAAGGTAAGGTGCTGTTTGTACCAACCTTGCACAACAGGTA	1200
Db	1513	GATCCTGAAGTTATGCGCAAGGTAAGGTGCTGTTTGTACCAACCTTGCACAATGCTGTA	1572
QY	1201	ACAGAAAGAAATTTTATGAAAAGTCAATTATGTCAGTTTGGGAAACTGGAAACGATTAAGAG	1260
Db	1573	ACAGAAAGAAATTTTATGAAAAGTCAATTATGTCAGTTTGGGAAACTGGAAACGATTAAGAG	1632
QY	1261	CTAAAGATTTATGCTTTCATTCATTTTATGATGAGAGATGAGTGTCAAGGCTATGGA	1320
Db	1633	TTTAAAGATTTATGCTTTCATTCATTTTATGATGAGAGATGAGTGTCTCAAGGCTATGGA	1692
QY	1321	GAATGAATGCTTAAAGCTTGAGGGGAGAAAAATTTGAATTTGTTTGTCTAAGCCACA	1380
Db	1693	GAATGAATGCTTAAAGCTTGAGGGGAGAAAAATTTGAATTTGTTTGTCTAAGCCACA	1752
QY	1381	GATCAGAGAGAGAAAGAAAGAAAGCTCAGAGGCAAGCAAGCAAGCAATATGATAT	1440
Db	1753	GATCAGAGAGAGAAAGAAAGAAAGCTCAGAGGCAAGCAAGCAAGCAATATGATAT	1812
QY	1441	GATTACTACTATTTATGTCACCTCATATGCTCCCTCCCAACAAGAGGTGCAAGGGGTGGA	1500
Db	1813	GATTACTACTATTTATGTCACCTCATATGCTCCCTCCCAACAAGAGGTGCAAGGGGTGGA	1872
QY	1501	GGTAGAGGTGGCTATGATATCTTCAGATTTATTTATGATACGAAGATTATATGATAT	1560
Db	1873	GGTAGAGGTGGCTATGATATCTTCAGATTTATTTATGATATGGAATATATATGATAT	1932
QY	1561	TATGTTATGATTAACATACTATCGTGGTGGATATGAAGATCCATACTATGGTTATGA	1620
Db	1933	TATGTTATGATTAACATACTATCGTGGTGGATATGAAGATCCATACTATGGTTATGA	1992
QY	1621	GATTTTCAAGTTTGAAGCTAGAGAGAAAGGCGGTATAGAGAGCAAGGGGTGCTGCTCATCC	1680
Db	1993	GATTTTCAAGTTTGAAGCTAGAGAGAAAGGCGGTATAGAGAGCAAGGGGTGCTGCTCATCC	2052
QY	1681	AGAGGTGCTGGGGGCTGCTCCCTCCCGGTGTAGAGCGGTTATTTACAAGAAGAGGGCCT	1740
Db	2053	AGAGGTGCTGGGGGCTGCTCCCTCCCGGTGTAGAGCGGTTATTTACAAGAAGAGGTCTT	2112
QY	1741	GGATCAGCAAGAGCGCTTCGCGGTGCGAGAGAGAGTGTGCCCAACAACAAGAGCGCGCGG	1800
Db	2113	GGATCAGCAAGAGCGGCTTCGAGGTGCGAGAGAGAGTGTGCCCAACAACAAGAGCGCGCGG	2172
QY	1801	GGAAAAAGGGGTGCAAGCGCGGTTC	1822
Db	2173	GTAAGTGGTGCAGAGGGGTGACC	2194

RESULT 10			
AF093821			
LOCUS	2373 bp	mRNA	linear
DEFINITION	Mus musculus RRM RNA binding protein GRV-RBP	mRNA,	complete cds.
ACCESSION	AF093821		
VERSION	AF093821.1	GI:3694985	
KEYWORDS			
SOURCE	Mus musculus.		
ORGANISM	Mus musculus		
	Eularyota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.		

REFERENCE 1 (bases 1 to 2373)  
 AUTHORS Du, G., Zhou, Y., Chen, J., Yuan, J., and Qiang, B.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-SEP-1998) Department of Biochemistry and Molecular  
 Biology, Institute of Basic Medical Sciences, 5 Dong Dan San Tiao,  
 Beijing 100005, P. R. China

FEATURES  
 Location/Qualifiers

1..2373  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 CDS  
 7..1884  
 /codon\_start=1  
 /product="RBM RNA binding protein GRV-RBP"  
 /protein\_id="AAC62511.1"  
 /db\_xref="GI:3694986"

/translation="MATEHYNGNGTEBEPDITTSATVHSNPTLLDAGLPQVAKELD  
 EIVAGLVASDDEKALKEPNDGALAVLOQKSDLSHVONKSAFLGVQKQY  
 RQREKGTVAADSKQDEPAKIKALERTGYLDVYTGKRGKPPDVSVGGQPSV  
 GEIEIVGKIPLDFEDELPLFEKAGPIMDLMDPLGLNKGYPFVETKEA  
 AOEAAKLNNHIEIRSGKHGVCSVANNELFVGSIPKSTKOILEEFSKTEGLTDV  
 ILYHODDDKKRGRPCFQEHVDHKAANOARRLMSKXVWNGVGVHEDADIEDDP  
 EYMAKVKVLPVRLNANTVTEHILEKSPSGPLERKVKIKDYAFTHFERDAVYAME  
 ENNGKDLBSENTEIVFAKPPDQREKRAKQKPAKQMDYIYTGPPHPPTRGRG  
 RGKRGVGPEDYGYEDYDYGYDYNHRYGVEDPYGYEDPQVGAHGRGARGG  
 AAPSRGGAAPPRGAGYSGRGAHPSGAVRGVGAQQRGSGVARGARSGANVGK  
 RADGYNQDPTKRQTNHQMWSQPIAQQLQSGDHSNHYGVKSEMEFYDPTFGQW  
 K"

BASE COUNT 740 a 412 c 587 g 634 t  
 ORIGIN

Query Match 46.6%; Score 1608; DB 10; Length 2373;  
 Best Local Similarity 98.1%; Pred. No. 0;

Matches 1652; Conservative 0; Mismatches 20; Indels 12; Gaps 2;

QY 148 GGAACATGCTACAGAACATGTTAATGAAATGTGCTGAAGAGCCCATGATACTACT 207  
 Db 1 GGAACATGCTACAGAACATGTTAATGAAATGTGCTGAAGAGCCCATGATACTACT 60  
 QY 208 TCAGCAGTTATCATTCAGAAATTTTTCAGACATTCGTTATCTGGTTTACACAGAAA 267  
 Db 61 TCAGCAGTTATCATTCAGAAATTTTTCAGACATTCGTTATCTGGTTTACACAGAAA 120  
 QY 268 GTTGTGAAATACTAGTAATTTTACGTTGACGGCTAGTTGCACTAGTATTTAGAT 327  
 Db 121 GTTGTGAAATACTAGTAATTTTACGTTGACGGCTAGTTGCACTAGTATTTAGAT 180  
 QY 328 GAAAGGCTATCGAAGCTTTAAAGAGTTCAATGAAGACGGCGATTGGCAGTGTCAA 387  
 Db 181 GAAAGGCTATCGAAGCTTTAAAGAGTTCAATGAAGACGGCGATTGGCAGTGTCAA 240  
 QY 388 CAGTTTAAAGACGATCTCTCATGTTTCAGAACAAAGTGCCTTTATATGTGAGTC 447  
 Db 241 CAGTTTAAAGACGATCTCTCTCATGTTTCAGAACAAAGTGCCTTTATATGTGAGTC 300  
 QY 448 ATGAAGACTTACAGGACAGAGAAAAACAGGGGACCAAGTAGACACTCTAGTAAAGA 507  
 Db 301 ATGAAGACTTACAGGACAGAGAAAAACAGGGGACCAAGTAGACACTCTAGTAAAGA 360  
 QY 508 CCAATGAGGCAAGATTAAAGCACTTTTGAAGAAACAGGCTTACACTTGAATGTGACT 567  
 Db 361 CCAATGAGGCAAGATTAAAGCACTTTTGAAGAAACAGGCTTACACTTGAATGTGACT 420  
 QY 568 ACAAGTCAGAGGAAGTATGAGAGACCACTCCAGATTCGTTTATTCAGGTACAGCCCT 627  
 Db 421 ACAAGTCAGAGGAAGTATGAGAGACCACTCCAGATTCGTTTATTCAGGTACAGCCCT 480  
 QY 628 TCTGTTGCACTGAGATATTGTGGGGAAGATCCCAAGATCTGTTTGAAGTGAAGCTT 687  
 Db 481 TCTGTTGCACTGAGATATTGTGGGGAAGATCCCAAGATCTGTTTGAAGTGAAGCTT 540  
 QY 688 GTTTCATTTATGAGAAAGCTGACCTATATGAGATCTTGTATTAATGATGATCCG--- 744  
 Db 541 GTTTCATTTATGAGAAAGCTGACCTATATGAGATCTTGTATTAATGATGATCCGCTC 600

QY 745 -----CTGACGTGCTCAACAGAGTTATGCTTGTGCACTTTTGTACAAAGAGCA 798  
 Db 601 ACTGCTCACTGCTCTCAATAGAGGTATAGCTTGTGCTACTTTTGTACAAAGAGCA 660  
 QY 799 GCACAGAGGCTGTTTAACTGATTAATATCATGAATAATCGTTCGGGAAGACATTTGT 858  
 Db 661 GCTCAGAGGCTGCTTAACCTGATTAATATCATGAATAATCGTTCGGGAAGACATTTGT 720  
 QY 859 GTCTGATCTCAGTTGCCAACATAGGCTTTTGTGGGCTGCAATTCCTAAGATAAAC 918  
 Db 721 GTCTGATCTCAGTTGCCAACATAGGCTTTTGTGGGCTGCAATTCCTAAGATAAAC 780  
 QY 919 AAGAGCAATCTTGAAGAAATTTAGCAAGAGGAGGCTCTACAGATGCTATTTA 978  
 Db 781 AAGAGCAATCTTGAAGAAATTTAGCAAGAGGAGGCTCTACAGATGCTATTTA 840  
 QY 979 TACCACTCACTGATGACAAAGAAAAAACAAGGCTTTTGTCTTGAATATGAAGAT 1038  
 Db 841 TACCACTCACTGATGACAAAGAAAAAACAAGGCTTTTGTCTTGAATATGAAGAT 900  
 QY 1039 CACAAAACGCTGCCAGGCAAGACGTAGGCTAATGATGTGTAAGTCAAAAGCTGGGA 1098  
 Db 901 CACAAAACGCTGCCAGGCAAGACGTAGGCTAATGATGTGTAAGTCAAAAGCTGGGA 960  
 QY 1099 AATGTTGAACTGTTGAGTGGGCTGATCTTATGAAGATCTGATCTGAAGTTATGCA 1158  
 Db 961 AATGTTGAACTGTTGAGTGGGCTGATCTTATGAAGATCTGATCTGAAGTTATGCA 1020  
 QY 1159 AAGGTAAAGTGTGTTTGTACGCAACTTGCACAACAGGTAACAGAAATTTTAGAA 1218  
 Db 1021 AAGGTAAAGTGTGTTTGTACGCAACTTGCACAACAGGTAACAGAAATTTTAGAA 1080  
 QY 1219 AAGTCAATTTAGTCAGTTTGGGAACTGAAACGAGTAAAGACTAAAGATTATGCTTC 1278  
 Db 1081 AAGTCAATTTAGTCAGTTTGGGAACTGAAACGAGTAAAGACTAAAGATTATGCTTC 1140  
 QY 1279 ATTCAATTTGATGAGAGATGCTGCTCAAGGCTATGGAAGAAATGATGTTAAGAC 1338  
 Db 1141 ATTCAATTTGATGAGAGATGCTGCTCAAGGCTATGGAAGAAATGATGTTAAGAC 1200  
 QY 1339 TTGAGGAGGAAATTTGAAATTTGTTTGTAAAGCCACAGATCAGAAAGGAAAGAA 1398  
 Db 1201 TTGAGGAGGAAATTTGAAATTTGTTTGTAAAGCCACAGATCAGAAAGGAAAGAA 1260  
 QY 1399 AGAAAAGCTCAGAGGCAAGCAAGAAATCAATGATGATGATTAATCTATTAATGCT 1458  
 Db 1261 AGAAAAGCTCAGAGGCAAGCAAGAAATCAATGATGATGATTAATCTATTAATGCT 1320  
 QY 1459 CCACTCATATGCTCTCCCAACAAAGAGTCAAGGCGTGAAGTGAAGTGGCTATGGA 1518  
 Db 1321 CCACTCATATGCTCTCCCAACAAAGAGTCAAGGCGTGAAGTGAAGTGGCTATGGA 1380  
 QY 1519 TATCCTCGAATTTATGATGATGAGAAATTAATGATTAATGATTAATGATTAATGAT 1578  
 Db 1381 TATCCTCGAATTTATGATGATGAGAAATTAATGATTAATGATTAATGATTAATGAT 1440  
 QY 1579 AACTATCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1638  
 Db 1441 AACTATCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
 QY 1639 AGAGGAAGGGGTGTGAGAGAGCAAGGGGTGCTCTCATCAAGAGTGTGGGGCTGCT 1698  
 Db 1501 AGAGGAAGGGGTGTGAGAGAGCAAGGGGTGCTCTCATCAAGAGTGTGGGGCTGCT 1560  
 QY 1699 CTTCCCGGTGTGAGAGCGGTTATTCAGAGAGAGCGCTGAGATCAGCAAGAGCGCTT 1758  
 Db 1561 CTTCCCGGTGTGAGAGCGGTTATTCAGAGAGAGCGCTGAGATCAGCAAGAGCGCTT 1620  
 QY 1759 CGCGGTGAGAGAGAGTCCCAACAAAGAGCGCGGAGGAGAAAGGCTGCAAGGCC 1818  
 Db 1621 CGCGGTGAGAGAGAGTCC---CAACAAAGAGCGCGGAGTACGTGTGCAAGGGGT 1677



QY 1534 TATGATACAGATTATTTATGATTATTTATGATTATTCACATACTATCGTGTGA 1593  
| | | | |  
Db 1381 TATGATATAGAAATTTATTTATGATTATTTATGATTATTCACATACTATCGTGTGA 1440  
| | | | |  
QY 1594 TATGATATACATCTATGTTATGAAATTTTCAAGTTGAGCTAGAGAGAGGTTGTT 1653  
| | | | |  
Db 1441 TATGATATACATCTATGTTATGAAATTTTCAAGTTGAGCTAGAGAGAGGTTGTT 1500  
| | | | |  
QY 1654 AGAGGAGCAAGGGGTCTGTCATCCAGAGGTCGTGGGGTCTCTCCCGGTGTGA 1713  
| | | | |  
Db 1501 AGAGGAGCAAGGGGTCTGTCATCCAGAGGTCGTGGGGTCTCTCCCGGTGTGA 1560  
| | | | |  
QY 1714 GCGGTTATTCAGAGAGAGGCGCTGATCAGCAAGAGCGCTGCGGTGTGAAGGA 1773  
| | | | |  
Db 1561 GCGGTTATTCAGAGAGAGGCGCTGATCAGCAAGAGCGCTGCGGTGTGAAGGA 1620  
| | | | |  
QY 1774 GGTGCCCAACAAAGAGGCGCGGGGAAAGGGTTCAGAGCGGTCTGACTGTTA 1833  
| | | | |  
Db 1621 GGTGCCCAACAAAGAGGCGCGGGGTAAGGGGTTCAGAGCGGTCTGACTGTTA 1680  
| | | | |  
QY 1834 CAT 1837  
| | | | |  
Db 1681 CAT 1684  
| | | | |  
RESULT 12  
AY034481 1872 bp mRNA linear PRI 03-OCT-2001  
LOCUS AY034481 Homo sapiens hnRNP Q3 mRNA, complete cds.  
DEFINITION AY034481  
ACCESSION AY034481  
VERSION AY034481.1 GI:15809585  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1872)  
Mourelatos, Z., Abel, L., Yong, J., Kataoka, N. and Drysdale, G.  
SMN interacts with a novel family of hnRNP and spliceosomal  
proteins  
JOURNAL EMBO J. 20 (19), 5443-5452 (2001)  
MEDLINE 21458434  
PUBMED 11574476  
REFERENCE 2 (bases 1 to 1872)  
Mourelatos, Z.  
AUTHORS Direct Submission  
TITLES Submitted (11-MAY-2001) HMMT/ Biochemistry, University of  
JOURNAL Pennsylvania, 330 CRB, 415 Curie Blvd., Philadelphia, PA 19104, USA  
FEATURES  
source location/Qualifiers  
1..1872  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1..1872  
/codon\_start=1  
/product="hnRNP Q3"  
/protein\_id="AAK59703.1"  
/db\_xref="GI:15809586"  
/translation="MATERHVNNGTEEPMDTTSVAIHSENFQTLDAIGLPOKVAEKLD  
EIVAGLVASDLDLDERALEALKPFNEDGALVLOPKSDLSHVONKSAFLGWNKY  
RORRKGTVASDSDKPEAKIKALLERGTLDVYTGKRYGPPDPDVSQGPSV  
GTEIFVYKIPDLPEDELPKPKAGPIVDLMDPLGLRGAFAVFCCKEAAOE  
AVKLVNHEIRSGKHGVCISVANNPLPGSI PKKTEQGLIERSKYTEGLTDVLY  
HOPDKKRSFCPLEYEDHKTAQARRRLMSGKRVMSNAGTIVEADAVAMEBN  
AKVAVLFVNLANVTYETLIERAFSOPGLERVKLKDVAFLHFERDAVAMEBN  
GDLGEGNEIIVFANRPDQKREKRAORQAANQWDDVYVYDPHMPPTGRGRG  
RGYGYPDYGYEDYDYDYDYDYNHRYGVEDPYGYEDPVGARGRGARGAAGP  
SGRGAAPRSGAGYSORSGPSAGRGARGAQOORGRGVARGAGGAGVGGKRG  
ADGVNODPSKRGQTNNQWGSQPIAQOPLQSGDHSNGYKSENGEPYDPTGCGMK"

BASE COUNT 590 a 323 c 489 g 470 t  
ORIGIN  
Query Match 45.0%; Score 1552.2; DB 9; Length 1872;  
Best Local Similarity 95.6%; Pred. No. 0;

Matches 1596; Conservative 0; Mismatches 73; Indels 0; Gaps 0;  
QY 154 ATGGCTACGAAACATGTTATATGAAATGGTACTGAAAGCCCATGATACCTTCAGCA 213  
| | | | |  
Db 1 ATGGCTACGAAACATGTTATATGAAATGGTACTGAAAGCCCATGATACCTTCAGCA 60  
| | | | |  
QY 214 GTTATCATTCGAAAAATTTTCAGACATGTCGTGATGCTGTTTCCAGAAAGTGTCT 273  
| | | | |  
Db 61 GTTATCATTCGAAAAATTTTCAGACATGTCGTGATGCTGTTTCCAGAAAGTGTCT 120  
| | | | |  
QY 274 GAAAACTAGATGAAATTTACGTTGCAAGGGCTAGTTGCAATAGTGAATTAATGAAAGA 333  
| | | | |  
Db 121 GAAAACTAGATGAAATTTACGTTGCAAGGGCTAGTTGCAATAGTGAATTAATGAAAGA 180  
| | | | |  
QY 334 GCTATGAAAGCTTTAAAGAGTCAATGAAAGCGCGCATGGCGCTTCAACAGTTT 393  
| | | | |  
Db 181 GCTATGAAAGCTTTAAAGAGTCAATGAAAGCGCGCATGGCGCTTCAACAGTTT 240  
| | | | |  
QY 394 AAAGACAGTATCTCTCATGTTGAGAAAGAAAGCTTTTATATGAGATCATGAG 453  
| | | | |  
Db 241 AAAGACAGTATCTCTCATGTTGAGAAAGAAAGCTTTTATATGAGATCATGAG 300  
| | | | |  
QY 454 ACTTACAGGCAAGAGAAAAACAGGGACCAAGTAGAGACCTTAGTAAAGAACAGAT 513  
| | | | |  
Db 301 ACTTACAGGCAAGAGAAAAACAGGGACCAAGTAGAGACCTTAGTAAAGAACAGAT 360  
| | | | |  
QY 514 GAGGCAAGATTATAGGCACTTTTGAAGAAAGAGCTACACCTTGATGTGACTACAGGT 573  
| | | | |  
Db 361 GAGGCAAGATTATAGGCACTTTTGAAGAAAGAGCTACACCTTGATGTGACTACAGGT 420  
| | | | |  
QY 574 CAGAGAGATATGAGAGACACCTCCAGATTCCTGTTATATGAGTACAGAGCTTCTGT 633  
| | | | |  
Db 421 CAGAGAGATATGAGAGACACCTCCAGATTCCTGTTATATGAGTACAGAGCTTCTGT 480  
| | | | |  
QY 634 GGCACATGATATTTGTGGGGAAGATCCCAAGATCTGTTAGAGATAGCTTGTCCA 693  
| | | | |  
Db 481 GGCACATGATATTTGTGGGGAAGATCCCAAGATCTGTTAGAGATAGCTTGTCCA 540  
| | | | |  
QY 694 TTATTTGAGAAAGCTGACCTTATATGAGATCTTCTGTTATATGAGATAGCTTGTCT 753  
| | | | |  
Db 541 TTATTTGAGAAAGCTGACCTTATATGAGATCTTCTGTTATATGAGATAGCTTGTCT 600  
| | | | |  
QY 754 CTCAACAGAGGTTAGCCCTTGTGCACTTTTGTACAAAAGAGCAGACAAAGCTGT 813  
| | | | |  
Db 601 CTCAACAGAGGTTAGCCCTTGTGCACTTTTGTACAAAAGAGCAGACAAAGCTGT 660  
| | | | |  
QY 814 AAACGTATATATATCATGAAATTCGTTCCGGGAAGACATGAGTGTGCACTTCAGTT 873  
| | | | |  
Db 661 AAACGTATATATATCATGAAATTCGTTCCGGGAAGACATGAGTGTGCACTTCAGTT 720  
| | | | |  
QY 874 GCGAACAATAGGCTTTTGTGGGCTCGATTCCTAAGATGAAACCAAGAGCAGATTC 933  
| | | | |  
Db 721 GCGAACAATAGGCTTTTGTGGGCTCGATTCCTAAGATGAAACCAAGAGCAGATTC 780  
| | | | |  
QY 934 GAGGAATTTAGCAAGATGACAGAGGCTCACAAGATGATTTTATACCACTGAT 993  
| | | | |  
Db 781 GAGGAATTTAGCAAGATGACAGAGGCTCACAAGATGATTTTATACCACTGAT 840  
| | | | |  
QY 994 GACAAAGAAAAAACAAGAGCTTTTGTCTTCTGATATGAGATCAAAAACAGCTGCC 1053  
| | | | |  
Db 841 GACAAAGAAAAAACAAGAGCTTTTGTCTTCTGATATGAGATCAAAAACAGCTGCC 900  
| | | | |  
QY 1054 CAGGCAAGAGCTAGCTAATGAGTGTAAAGTCAAAAGCTGAGGAAATTTGAAACTGTT 1113  
| | | | |  
Db 901 CAGGCAAGAGCTAGCTAATGAGTGTAAAGTCAAAAGCTGAGGAAATTTGAAACTGTT 960  
| | | | |  
QY 1114 GAGTGGGCTGATCTATGAAAGATCTGATTCCTGAAAGTATAGCAAGATTAAGTGTG 1173  
| | | | |  
Db 961 GAGTGGGCTGATCTATGAAAGATCTGATTCCTGAAAGTATAGCAAGATTAAGTGTG 1020  
| | | | |  
QY 1174 TTGTATAGCAACCTTGCCAAACAGGTAAACAGAAAGAAATTTGAAAGATCATTTAGTCA 1233  
| | | | |  
Db 1021 TTGTATAGCAACCTTGCCAAACAGGTAAACAGAAAGAAATTTGAAAGATCATTTAGTCA 1080  
| | | | |

Query Match	Best Local Similarity	Score	DB	Length
Matches 1464; Conservative	99.2%; Pred. No. 8.5e-305;	0; Mismatches 12; Indels 0; Gaps 0;		
BASE COUNT	689 a	387 c	551 g	605 t
ORIGIN				
Query Match	42.2%; Score 1456.8; DB 10; Length 2232;			
1 TAAAGAGTTCATGAGAACGGCGCATTTGGCAGTGTCTTCAACAGTTTAAAGACAGTATC				
Db	1 TAAAGAGTTCATGAGAACGGCGCATTTGGCAGTGTCTTCAACAGTTTAAAGACAGTATC	60		
Qy	347 TAAAGAGTTCATGAGAACGGCGCATTTGGCAGTGTCTTCAACAGTTTAAAGACAGTATC	406		
Db	1 TAAAGAGTTCATGAGAACGGCGCATTTGGCAGTGTCTTCAACAGTTTAAAGACAGTATC	60		
Qy	407 TCTCTCATGTTCAAGAACAAAGTGCGCTTTATATGTGAGTCAATGAACCTTACAGGCGCA	466		
Db	61 TCTCTCATGTTCAAGAACAAAGTGCGCTTTATATGTGAGTCAATGAACCTTACAGGCGCA	120		
Qy	467 GAGAAAAACAGGGGACCAAAATGACAGACTTAAGTAAAGACAGATAGGCGCAAGATTA	526		
Db	121 GAGAAAAACAGGGGACCAAAATGACAGACTTAAGTAAAGACAGATAGGCGCAAGATTA	180		
Qy	527 AGGCACTTTGGAAAAAGAACAGGCTACACACTTGTATGACTACAGGTCAAGAGAAATAG	586		
Db	181 AGGCACTTTGGAAAAAGAACAGGCTACACACTTGTATGACTACAGGTCAAGAGAAATAG	240		
Qy	587 GAGGACCACTTCAGATTCGGCTTATTAGGTACAGACGCTTCTGTTGGGACTGAGATAT	646		
Db	241 GAGGACCACTTCAGATTCGGCTTATTAGGTACAGACGCTTCTGTTGGGACTGAGATAT	300		
Qy	647 TTGTGGGGAAGATCCCGAGAGATCTGTTTGAAGATGAGACTTGTTCATTATTGGAAG	706		
Db	301 TTGTGGGGAAGATCCCGAGAGATCTGTTTGAAGATGAGACTTGTTCATTATTGGAAG	360		
Qy	707 CTGACCTTATATGGATTTTGTTTAATGATGATCCGCTCACTGTGTCAACAGAGTT	766		
Db	361 CTGACCTTATATGGATTTTGTTTAATGATGATCCGCTCACTGTGTCAACAGAGTT	420		
Qy	767 ATGCGTTGTCACTTTTGTCAAAAAGAGACAGACAAGAGGCTGTGTAACCTGTATAATA	826		
Db	421 ATGCGTTGTCACTTTTGTCAAAAAGAGAGACAGACAAGAGGCTGTGTAACCTGTATAATA	480		
Qy	827 ATCATGAATTCGTTCCGGGAAGACATTTGTTGTGATCTCAAGTTGCCAACATAGGC	886		
Db	481 ATCATGAATTCGTTCCGGGAAGACATTTGTTGTGATCTCAAGTTGCCAACATAGGC	540		
Qy	887 TTTTGTGGGCGCATTCCTTAAGATTAACAAGAGGACAGATCTTGAGGAAATTTAGCA	946		
Db	541 TTTTGTGGGCGCATTCCTTAAGATTAACAAGAGGACAGATCTTGAGGAAATTTAGCA	600		
Qy	947 AAGTACAGAGGGTCTCAAGATGTCAATTTTATACACACACCTGATGACAAAGAAAAA	1006		
Db	601 AAGTACAGAGGGTCTCAAGATGTCAATTTTATACACACACCTGATGACAAAGAAAAA	660		
Qy	1007 ACAAGGCTTTTGTCTTTTGAATATGAAGTCACAAAAACAGCTGCCACAGCAGAGCTA	1066		
Db	661 ACAAGGCTTTTGTCTTTTGAATATGAAGTCACAAAAACAGCTGCCACAGCAGAGCTA	720		
Qy	1067 GCGTATATGATGTATAAGTCAAAAGTCTGGGGAAGATGTGAACTGTGAGGGCTGATC	1126		
Db	721 GCGTATATGATGTATAAGTCAAAAGTCTGGGGAAGATGTGAACTGTGAGGGCTGATC	780		
Qy	1127 CTATTGAAGATCTGATCTGAAGTTATGCAAAAGTAAAGTAAAGTCTGTTGTACGCAACC	1186		





Query	Best Local Match	Similarity	Score	DB %	Length	Mismatches	Indels	Gaps
1800	GGGAAAAGGGGTCAGAGCGCGGTCCTGACCTGTATCAATGAAGACGACTGCTATTTGG	39.3%	1356.4	2039	74337	41	24	14
28323	GGGAAAAGGGGTCAGAGCGCGGTCCTGACCTGTATCAATGAAGACGACTGCTA-TGGG	36.0%	1263.7	2039	74337	41	24	14
1860	GATTACACCAAGAGCTTCAGTGAAGTATGTAAGAAATCAAGCAACCTTAATATC	39.3%	1356.4	2039	74337	41	24	14
28264	GATTACACCAAGAGCTTCAGTGAAGTATGTAAGAAATCAAGCAACCTTAATATG	36.0%	1263.7	2039	74337	41	24	14
1920	TGCGATGTAAGAGCATATTCATTTGCAAGAACCTTCCTATGAAGTCATGAATCA	39.3%	1356.4	2039	74337	41	24	14
28205	TGCGATGTAAGAGCATATTCATTTGCAAGAACCTTCCTATGAAGTCATGAATCA	36.0%	1263.7	2039	74337	41	24	14
1980	ATACGGGACATTTGAACCTTAATTTGTAATGTAATTTCTTAACAATTTCTCG	39.3%	1356.4	2039	74337	41	24	14
28145	ATACGGGACATTTGAACCTTAATTTGTAATGTAATTTCTTAACAATTTCTCG	36.0%	1263.7	2039	74337	41	24	14
2040	CAGTCAGATTAATTAAGCTAATTTTCCAAATGTGTCCA-AAAAATCCT	39.3%	1356.4	2039	74337	41	24	14
28085	CAGTCAGATTAATTAAGCTAATTTTCCAAATGTGTCCA-AAAAATCCT	36.0%	1263.7	2039	74337	41	24	14
2099	TCATTAACCTTAAGCTATGTAATTAATTAAGAAATGTTGCTTTTAAATAATCGC	39.3%	1356.4	2039	74337	41	24	14
28025	TCATTAACCTTAAGCTATGTAATTAATTAAGAAATGTTGCTTTTAAATAATCGC	36.0%	1263.7	2039	74337	41	24	14
2159	TCTAAGTGAATTTTCCCTCTTTTAAATTAAGATTCAGAGTGTATTTGAATA	39.3%	1356.4	2039	74337	41	24	14
27970	TCTAAGTGAATTTTCCCTCTTTTAAATTAAGATTCAGAGTGTATTTGAATA	36.0%	1263.7	2039	74337	41	24	14
2219	TTCTCTGAATTTGTGCAATTAATTTTATGCAATGATCAAG-ATGCCATGTTGG	39.3%	1356.4	2039	74337	41	24	14
27911	TTCTCTGAATTTGTGCAATTAATTTTATGCAATGATCAAG-ATGCCATGTTGG	36.0%	1263.7	2039	74337	41	24	14
2275	TACCTTAATTTTATTTCTGCTCAGCAAGGTATATCATATTTGTATATCTTTTAA	39.3%	1356.4	2039	74337	41	24	14
27851	TATCTTTAA-TTTTATTTGCTCAGCAAGGTATATCATATTTGTATATCTTTTAA	36.0%	1263.7	2039	74337	41	24	14
2335	TAGTAATCACTTTGAATTTGTGTCAG-ATGCACTTTCAAGTGTATATCATGAGCTGG	39.3%	1356.4	2039	74337	41	24	14



Db	27792	TAGTATCACTTTTGAATTTGTGTCAGATATGAGTTTCAGGTTAAATCATCAGAGCTGG	27735
Qy	2393	TTAAGTCAGGCATTCGAGATATAGTGGTCTTTTCAGAACCTTTT-AAAGGTTGGTTAACT	2451
Db	27732	TTAAGTCAGGCATTCGAGATATAGTGGTCTTTTCAGAACCTTTTAAAGAGGTTGGTTAACT	26739
Qy	2452	ACCTCAGTGCAGAGAGATTGAACATAACCTGTCTGTATCTGTACATATAGAAAAATCTTTGTA	2511
Db	27672	ACCTCAGTGCAGAGAGATTGAACATAACCTGTCTGTATCTGTACATATAGAAAAATCTTTGTA	27613
Qy	2512	GATTAAGCAAGGCTTGTTAAATATGATATGAGGGTAAAGATTTTAATATACAAATGTAA	2571
Db	27612	GATTAAGCAAGGCTTGTTAAATATATATATGAGGGTAAAGATTTTAATATACAAATGTAA	27553
Qy	2572	CATCTTACTGTGCTTTACTATTTTCAGAGGCTTGTAACATCTTCTCATGACATCATACAG	2631
Db	27552	CATCTTACTGTGCTTTACTATTTTCAGAGGCTTGTAACATCTTCTCATGACATCATACAG	27493
Qy	2632	GCCCTGCTTTTTCGTAATTTTGTGCTGTAAGAAAGCAGCCTTGCTCTTCAGATATTGTAG	2691
Db	27492	GCCCTGCTTTTTCGTAATTTTGTGCTGTAAGAAAGCAGCCTTGCTCTTCAGATATTGTAG	27433
Qy	2692	TTATTTGGATGTATATATAGTTTATGCAAGATGTTACTTTTGTAAAGACATCAAGATGTTCAA	2751
Db	27432	TTATTTGGATGTATATATAGTTTATGCAAGATGTTACTTTTGTAAAGACATCAAGATGTTCC	27376
Qy	2752	AAAAAAGTGCATCCGAACTTTGATCTTAATAATACGAGTGCCTTTATATAAAAGTCAGCT	2811
Db	27375	AAAAAAGTGCATCCGAACTTTGATCTTAATAATACGAGTGCCTTTATATAAAAGTCAGCT	27316
Qy	2812	AAACTGCAATTTGTACAGCAAAAGCCTGCATTTGGATATTTTGAAGTTTTCATTAAT	2871
Db	27315	AAACTGCAATTTGTATACAGCAAGCCTGCATTTGGATATTTTGAAGTTTTCATTAAT	27256
Qy	2872	CATTAATAATTATATATAGTCTGTATGTTTACCTTTTATAGTAAAGGATGTTTCATTAAT	2931
Db	27255	CATTAATAATTATATATAGTCTGTATGTTTACCTTTTATAGTAAAGGATGTTTCATTAAT	27196
Qy	2932	GCAATTTGTTATTTGCTGATCACTATATAAAATGTGAATCAGCTTTCATTTCTTATCAGAGT	2991
Db	27195	GCAATTTGTTTCCGTGCTGATCACTGTAACATGTGAATCAGCTTTCATTTCTTATCAGAGT	27136
Qy	2992	CATGATTAATCTGTATATATAGTATCAGTACATATTTTGCTATATGTTTTTAATTTCTAAAGCA	3051
Db	27135	CATGATTAATCTGTATATATAGTATCAGTACATATTTTGCTATATGTTTTTAATTTCTAAAGCA	27076
Qy	3052	CCTTATATGACATATGATGTTCAGTGTGAAGATCCTCATTTGATATCACCCCTCAAAAAAT	3111
Db	27075	CCTTATATGACATGATGTTCAGTGTGAAGATCCTCATTTGATATCACCCCTC-AAAAAT	27017
Qy	3112	TTTTTTGCGAAGTCTTAAGTTGATATAGCTTAAAGTCAAAAGT-AAAAATTATGTTTAAAGTA	3170
Db	27016	TTTTTTGCGAAGTCTTAAGTTGATATAGCTTAAAGTCAAAAGTATTTATAGTTTCATTTA	26957
Qy	3171	GGAATTTGGTGTAAAGAAACACCTT-CCCCCTTCCCAAAAGGATATGCGAGTTCTATCAC	3229
Db	26956	GGAATTTGGTGTAAAGAAATCCCTTCCCCCTTCCCAAAAGGATATGCGAGTTATATCAC	26897
Qy	3230	ACAACCCAGTAGGSCACAAGATACGATCAGAC-TAGACCTGATCAAGGTTTATATACACA	3288
Db	26896	ATACCCATATAGGSCACAAGATATAGATCAGACCTTATACTTTATTTAAGGTTTATATACACA	26837
Qy	3289	CCAGTTTCCCGAGTAATGTCAAAATTTTAAACAAGAAATTTAGACATGTCAATATGTTCAAAATG	3348
Db	26836	CCAGTTTCCCGAGTAATGTCAAAATTTTAAACAAGAAATCAGACATGTCAATATGTTCAAAATG	26777
Qy	3349	CTCATGGGCAAAACATCATTTTGCAATTCCTCGCAAAATATTAATTTGTTTATAC	3398
Db	26776	CTCATGGGCAAAACATCATTTTGCAATTCCTCGCAAAATATTAATTTGTTTATAC	26727

LOCUS	AY034482	1767 bp	mRNA	linear	PRI 03-OCT-2001
DEFINITION	Homo sapiens hnRNP Q2 mRNA, complete cds.				
ACCESSION	AY034482				
VERSION	AY034482.1	GI:15809587			
KEYWORDS					
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 1767)				
TITLE	Mourtelatos,Z., Abel,L., Yong,J., Kataoka,N. and Dreyfuss,G.				
JOURNAL	SNM interacts with a novel family of hnRNP and spliceosomal proteins				
FEATURES	EMBO J 20 (19), 5443-5452 (2001)				
source	1..1767				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	1..1767				
	/codon_start=1				
	/product="hnRNP Q2"				
	/protein_id="AAK59704.1"				
	/db_xref="GI:15809588"				
	/translation="MATEBHNGNTEPNTTSAVHSENPQTLLDAGLFQKYAEKID				
	EIVYAGLVASHSDDERAIEALKEFNEDGALAVIQFQSDLSHVQNKSAPLQGVWKY				
	RQREKQGTKVADSSKGPDEAKIKALERTYTLDDVTGKRGGPPDSVYSGQSPV				
	GTELFPVKIKIPDLFDEDELVPLFEKAGIPMLPLMDPLTLGNGYAVFTCKEAAOE				
	AVKLNNHRIIRSGKILGVCI SVANNRLRFVSIIPSKKEOILIEFSVTEGTLIDVILY				
	HOPDKKNSRSGFGLYEYDEHDKTAQNVPLVRNALNTVEIIEKASQSGKLERVKK				
	LKVQAPLHPDERDSAVYAMEMNMGKDLSEGNIEIVAKPDPQKREKKAQROAKNOM				
	YDDIYTYGPHMPPETKRGARGGAGGAGYGPDDYGYEDYDYGYDHNHNGGYEDPY				
	YGVDELFGVARGRGGRGARGAPRGRGAPRGRGYSORGGPGRGARGVARGAGGAG				
	QQRORVARGARGGAGGARGKRGADGYNPDPSRRQTNQNMGSPIAQOPLQGDHS				
	GNYGYKSENOEFPVDYTGQOQWK"				
BASE COUNT	561 a	310 c	453 g	443 t	
ORIGIN					
Query Match	38.9%	Score 1343.4	DB 9	Length 1767	
Best Local Similarity	89.8%	Pred. No. 3e-280			
Matches 1498	Conservative 0	Mismatches 66	Indels 105	Gaps 1	
QY	154	ATGGCTACGAAACATGTTAATGGAATGTACTGAAGAGCCCATGATGACTACTTCACGA	213		
DB	1	ATGGCTACGAAACATGTTAATGGAATGTACTGAAGAGCCCATGATGACTACTTCGCA	60		
QY	214	GTATTCATTGAGAAATTTTTCAGACATGCTGATGCTGGTTTACACAGAAAGTTGCT	273		
DB	61	GTATTCATTGAGAAATTTTTCAGACATGCTGATGCTGGTTTACACAGAAAGTTGCT	120		
QY	274	GAAGAACTAGATGAATTTACGTTGACGGGCTAGTTGCACTAGATGATTGATGAAGA	333		
DB	121	GAAGAACTAGATGAATTTACGTTGACGGGCTAGTTGCACTAGATGATTGATGAAGA	180		
QY	334	GCTATCGAAGCTTTAAAGAGTTCAATGAGACGGCGCATTTGGCAGTGCCTTCAACAGTTT	393		
DB	181	GCTATCGAAGCTTTAAAGAGTTCAATGAGACGGCGCATTTGGCAGTGCCTTCAACAGTTT	240		
QY	394	AAAGCAGTGATCTCTCCATGCTCAGAAACAAAGTGCCCTTTTAATGTGAGTCAATGAAG	453		
DB	241	AAAGCAGTGATCTCTCCATGCTCAGAAACAAAGTGCCCTTTTAATGTGAGTCAATGAAG	300		
QY	454	ACTTACAGGACAGAT	513		
DB	301	ACTTACAGGACAGAT	360		
QY	514	GAGCGAAGATTTAAGGACCTTTGGAAGAGACAGGCTTACACACTTGTAGTGAATCAAGT	573		

```

Db 361 GAGGCAAAATTTAGGCACTTTGGAAAGACAGGCTACACATGTAATGACACCTGSA 420
Qy 574 CAGAGAGATATGAGAGACCACTCCAGATTCGGTTATTCAGGTGACAGCCCTTGTT 633
Db 421 CAGAGAGATATGAGAGACCACTCCAGATTCGGTTATTCAGGTGACAGCCCTTGTT 480
Qy 634 GGCACATGATATTTGGGGGAGAAATCCCAAGATCTGTTGAGATGAGCTTGTTCCA 693
Db 481 GGCACATGATATTTGGGGGAGAAATCCCAAGATCTGTTGAGATGAGCTTGTTCCA 540
Qy 694 TTAATTTAGAAAGCTGACCTATATGGAATCTTGTTAATGATGATGACGCTGCTGTT 753
Db 541 TTAATTTAGAAAGCTGACCTATATGGAATCTTGTTAATGATGATGACGCTGCTGTT 600
Qy 754 CTCAACAGAGTTATGCGTTTGTCACTTTTGTACAAAGAGCAGCAAGAGGCTGTT 813
Db 601 CTCAATAGAGTTATGCGTTTGTCACTTTTGTACAAAGAGCAGCTCAGAGGCTGTT 660
Qy 814 AAACGTATATATATCATGAAATTCGTTCCGGAGACATTTGGTCTGATCTCAGTT 873
Db 661 AAACGTATATATATCATGAAATTCGTTCCGGAGAACATATTTGGTCTGATCTCAGTT 720
Qy 874 GCCAACAATAGGCTTTTGTGGGCTCGATTCCTAAGATGAAACCAAGAGCAGATCTT 933
Db 721 GCCAACAATAGGCTTTTGTGGGCTCGATTCCTAAGATGAAACCAAGAGCAGATCTT 780
Qy 934 GAGGAATTTAGCAAGATGACAGAGGCTCAAGATGCTATTTATACCAACCACTGAT 993
Db 781 GAGGAATTTAGCAAGATGACAGAGGCTCAAGATGCTATTTATACCAACCACTGAT 840
Qy 994 GACAGAGAAATTAACAGAGGCTTTGCTTCTGTAATATGAAATGCAAAACAGCTGCC 1053
Db 841 GACAGAGAAATTAACAGAGGCTTTGCTTCTGTAATATGAAATGCAAAACAGCTGCC 900
Qy 1054 CAGGCAAGACGTAGCTATATGATGCTAAAGCTGAGGAAATGTTGAACTGTT 1113
Db 901 C----- 901
Qy 1114 GAGTGGCTGATCCTATTGAAGATCTGATCTGAAGTTATGGCAAGATAAAGTGTG 1173
Db 902 -----AGTAAAGTGTG 915
Qy 1174 TTTGTAGCAACCTTGCCACACAGGTAACAGAGAAATTTAGAAAGTCAATTAAGTCAG 1233
Db 916 TTTGTAGCAACCTTGCCACACAGGTAACAGAGAAATTTAGAAAGTCAATTAAGTCAG 975
Qy 1234 TTTGGAAACTGAAACGATGAAAGAGCTAAAGATTAAGCTTTCATTCAATTTGATGAG 1293
Db 976 TTTGGAAACTGAAACGATGAAAGAGTAAAGATTAAGCTTTCATTCAATTTGATGAG 1035
Qy 1294 AGAGATGCTGCTCAAGGCTATGAGAGAAATGAATGCTAAAGACTTGAGAGAGAAAT 1353
Db 1036 CGAGATGCTGCTCAAGGCTATGAGAGAAATGAATGCTAAAGACTTGAGAGAGAAAT 1095
Qy 1354 ATTGAATTTGTTTGTCTAACCACAGATCAAGAGAGAGAAAGAAAGCTCAGAGG 1413
Db 1096 ATTGAATTTGTTTGTCTAACCACAGATCAAGAGAGAGAAAGAAAGCTCAGAGG 1155
Qy 1414 CAAGCAGCAAGAAATCAAAATGATGATGATTAATTAATGCTCACTCATATGCTT 1473
Db 1156 CAAGCAGCAAGAAATCAAAATGATGATGATTAATTAATGCTCACTCATATGCTT 1215
Qy 1474 CCCCCAACAAGAGTGAAGGCGTGAAGGTAAGAGTGCATATGATATCTTCAGATTAAT 1533
Db 1216 CCTCCAAACAAGAGTGAAGGCGTGAAGGTAAGAGTGCATATGATATCTTCAGATTAAT 1275
Qy 1534 TATGATATCGAAGATTAATTAATGATGATGATTAATTAATGATTAATTAATGATGATG 1593
Db 1276 TATGATATGAGATTAATTAATGATGATGATTAATTAATGATTAATTAATGATGATG 1335
Qy 1594 TATGAAGATCCATATGATGATTAATGAAGATTTCAAGTTGAGCTAGAGAGAGGCTG 1653

```

```

Db 1336 TATGAAGATCCATATGATGATTAATGAAGATTTCAAGTTGAGCTAGAGAGGCTG 1395
Qy 1654 AGAGAGCAAGGGGCTGCTGCTCATCCAGAGGCTGTTGGGCTGCTCTCCCGTGATGA 1713
Db 1396 AGAGAGCAAGGGGCTGCTGCTCATCCAGAGGCTGTTGGGCTGCTCTCCCGTGATGA 1455
Qy 1714 GCGGTTATTTCAAGAGAGAGGCGCTGATCAGCAAGAGGCGTTGGCGTGAGAGAGA 1773
Db 1456 GCGGTTATTTCAAGAGAGAGGCGCTGATCAGCAAGAGGCGTTGGCGTGAGAGAGA 1515
Qy 1774 GGTGCCCAACAAACAAAGAGGCGCGGGGAGAAAGGGGCTGAGGCGGCTC 1822
Db 1516 GGTGCCCAACAAACAAAGAGGCGCGGGGATACGTGTTGCGAGAGGCTG 1564

```

Search completed: April 28, 2003, 00:54:54  
Job time : 18678.6 secs

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: April 27, 2003, 16:49:41 ; Search time 8469.41 Seconds

(without alignments)  
5783.170 Million cell updates/sec

Title: US-09-821-687-3\_COPY\_154\_1836

Perfect score: 1683

Sequence: 1 atggtctacgacatctttaa.....ccggtccctgactgttaca 1683

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_ses:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_ses:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_ocher:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pin:\*  
35: em\_htg\_mam:\*  
36: em\_htg\_rtd:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_higo\_hum:\*  
40: em\_higo\_mus:\*  
41: em\_higo\_ocher:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1683	100.0	3452	10	AB035725 Mus muscu
2	1602	95.2	2373	10	AF093821 Mus muscu
3	1583.8	94.1	1686	9	AY034483 Homo sapi
4	1574	93.5	2221	9	AF155568 Homo sapi
5	1552.2	92.2	1872	9	AY034481 Homo sapi
6	1552.2	92.2	2932	9	AF037448 Homo sapi
7	1519	90.3	2079	6	AR177851 Homo sapi
8	1456.8	86.6	2232	10	AF408434 Mus muscu
9	1442.8	85.7	103819	9	HSJ775C13 Human DNA
10	1442.8	85.7	212659	2	AL390737 Homo sapi
11	1366.8	81.2	3319	9	BC032643 Homo sapi
12	1343.4	79.8	1767	9	AY034482 Homo sapi
13	1222.8	72.7	3101	6	AX011753 Sequence
14	896	53.2	1899	10	AF441128 Mus muscu
15	895.8	53.2	151297	2	AC118904 Rattus no
16	875	52.0	2644	9	AF000364 Homo sapi
17	860.8	51.1	221309	2	AC122217 Mus muscu
18	856	50.9	2722	9	BC001449 Homo sapi
19	853.8	50.7	2471	10	BC004679 Mus muscu
20	816	48.5	181810	2	AC126148 Rattus no
21	762.4	45.3	121304	2	AC079134 Mus muscu
22	762.4	45.3	196488	2	AC084410 Mus muscu
23	736.4	43.8	934	6	AX013715 Sequence
24	706.2	42.0	202374	10	AC122868 Mus muscu
25	636.8	37.8	139996	9	AL161799 Human DNA
26	570.6	33.9	129237	2	AC127932 Rattus no
27	484	28.8	110783	2	AC129693 Rattus no
28	426	25.3	118429	2	AC119124 Rattus no
29	398.4	23.7	57824	2	AC101053 Mus muscu
30	381.8	22.7	2606	3	AY058477 Drosophila
31	368	21.9	384	6	AR079026 Sequence
32	358.2	21.3	74337	9	AL136082 Human DNA
33	354	21.0	2195	10	BC004001 Mus muscu
34	351	20.9	175221	2	AC11832 Rattus no
35	289.4	17.2	2467	9	BC028588 Homo sapi
36	250.8	14.9	1785	10	AY028945 Rattus no
37	250.8	14.9	1863	10	AF290984 Rattus no
38	242.4	14.4	1770	9	AJ272079 Homo sapi
39	242.4	14.4	1794	9	AF272078 Homo sapi
40	242.4	14.4	1945	9	AF209192 Homo sapi
41	240.8	14.3	1761	9	AF271789 Homo sapi
42	223	13.8	1911	9	AF271790 Homo sapi
43	228.6	13.6	117535	10	AL672208 Mouse DNA
44	228.6	13.6	218193	2	AC121608 Mus muscu
45	222.8	13.2	2105	9	AK000324 Homo sapi

#### ALIGNMENTS

RESULT 1  
AB035725  
LOCUS AB035725 3452 bp mRNA linear ROD 31-MAR-2000  
DEFINITION Mus musculus SYNCRIP mRNA, complete cds.  
ACCESSION AB035725  
VERSION AB035725.1 GI:6576814  
KEYWORDS SYNCRIP.  
SOURCE Mus musculus (strain:BDV) postnatal 21 day cDNA to mRNA.  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (sites) Mizutani, A., Fukuda, M., Iibata, K., Shitashiki, Y. and Mikoshiba, K.  
SYNCRIP, a cytoplasmic counterpart of heterogeneous nuclear ribonucleoprotein R, interacts with ubiquitous synaptotagmin

Pred. No. is the number of results predicted by chance to have a

JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

isoforms  
J. Biol. Chem. 275 (13), 9823-9831 (2000)  
20200483  
2 (bases 1 to 3452)  
Mizutani, A., Fukuda, M. and M. Koshihara, K.  
Direct Submission  
Submitted (11-DEC-1999) Akihiro Mizutani, Brain Science Institute,  
RIKEN, the Laboratory for Developmental Neurobiology, 2-1 Hirosawa,  
Wako, Saitama 351-0198, Japan (E-mail: amizutani@ims.u-tokyo.ac.jp,  
Tel: +81-48-467-9745, Fax: +81-48-467-9744)

## FEATURES

## Source

1. 3452  
/organism="Mus musculus"  
/strain="DDY"

## gene

/db\_xref="taxon:10090"  
/dev\_stage="postnatal 21 day"

## mRNA

1. 3452  
/gene="SYNCRIP"

## CDS

/evidence="experimental"  
154..1839  
/gene="SYNCRIP"

/protein\_id="BA68342.1"

/db\_xref="gi:6576815"

BASE COUNT 1056 a 623 c 777 g 996 t  
ORIGIN

Query Match 100.0%; Score 1683; DB 10; Length 3452;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1683; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGCTACAGAACTGTAATGAAATGTAAGGAGCCCATGAGTACTTACGCA 60  
154 ATGGCTACAGAACTGTAATGAAATGTAAGGAGCCCATGAGTACTTACGCA 213  
61 GTTATCATTCAGAAATTTTCAACATTCGTTGCTGTTTACACAGAAAGTGGT 120  
214 GTTATCATTCAGAAATTTTCAACATTCGTTGCTGTTTACACAGAAAGTGGT 273  
121 GAAAGAACTGTAATGTAATGTAAGGAGGCTAGTGCACATGATGATTGATGAAAGA 180  
274 GAAAGAACTGTAATGTAATGTAAGGAGGCTAGTGCACATGATGATTGATGAAAGA 333  
181 GCTATCGAAGCTTTAAAGAGTTCAATGAGCGGCGATTGGCAGTCTCAACAGTTT 240  
334 GCTATCGAAGCTTTAAAGAGTTCAATGAGCGGCGATTGGCAGTCTCAACAGTTT 393  
241 AAAGACAGTGTCTCTCATGTTCAAGCAAAAGTGCCTTTTATGAGATCATGAAG 300  
394 AAAGACAGTGTCTCTCATGTTCAAGCAAAAGTGCCTTTTATGAGATCATGAAG 453  
301 ACTTACAGGAT 360  
454 ACTTACAGGAT 513  
361 GAGGCAAGATTTAAGGAGCTTTTGAAGAGAGAGGCTACACTTGAATGATGATGAGAT 420  
514 GAGGCAAGATTTAAGGAGCTTTTGAAGAGAGAGGCTACACTTGAATGATGATGAGAT 573  
421 CAGGCAAGATTTAAGGAGCTTTTGAAGAGAGAGGCTACACTTGAATGATGATGAGAT 480

Db 574 CAGAGAGATGAT 633  
Qy 481 GGCACAGAGATTTTGGGAT 540  
Db 634 GGCACAGAGATTTTGGGAT 693  
Qy 541 TTATTTAGAT 600  
Db 694 TTATTTAGAT 753  
Qy 601 CTCACAGAGAGATTTGAT 660  
Db 754 CTCACAGAGAGATTTGAT 813  
Qy 661 AAACGTAT 720  
Db 814 AAACGTAT 873  
Qy 721 GCCAAGATAGGCTTTTGGGAT 780  
Db 874 GCCAAGATAGGCTTTTGGGAT 933  
Qy 781 GAGGATTTAGCAAGAT 840  
Db 934 GAGGATTTAGCAAGAT 993  
Qy 841 GACAGAT 900  
Db 994 GACAGAT 1053  
Qy 901 CAGGCAAGAT 960  
Db 1054 CAGGCAAGAT 1113  
Qy 961 GAGTGGCTGATCTTATGAT 1020  
Db 1114 GAGTGGCTGATCTTATGAT 1173  
Qy 1021 TTTGATGCAAGCTTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1080  
Db 1174 TTTGATGCAAGCTTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1233  
Qy 1081 TTTGGAAGAT 1140  
Db 1234 TTTGGAAGAT 1293  
Qy 1141 AGAGATGCTGCTCAAGAT 1200  
Db 1294 AGAGATGCTGCTCAAGAT 1353  
Qy 1201 ATTGAATTTGTTTGTATGAT 1260  
Db 1354 ATTGAATTTGTTTGTATGAT 1413  
Qy 1261 CAAGCAGAT 1320  
Db 1414 CAAGCAGAT 1473  
Qy 1321 CCCCACAGAT 1380  
Db 1474 CCCCACAGAT 1533  
Qy 1381 TATGATACAGAGATTTATGATTTATGATTTATGATTTATGATTTATGATTTATGATTTAT 1440  
Db 1534 TATGATACAGAGATTTATGATTTATGATTTATGATTTATGATTTATGATTTATGATTTAT 1593  
Qy 1441 TATGAGATCCATATATATGATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1500  
Db 1594 TATGAGATCCATATATGATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1553  
Qy 1501 AGAT 1560  
Db 1654 AGAT 1713

QY	1561	GGCGGTATTATTCACAGGAGGAGCCCTTGATCAGCAAGAAGGCGTTTGC CGGTGCGAAGA	1620
Db	1714	GCCGGTTATTATTCACAGGAGGAGCCCTTGATCAGCAAGAAGGCGTTTGC CGGTGCGAAGA	1773
QY	1621	GGTGCCCAACAACAAGAGGCGCGCGGGGAAAAAGGGGTGAAGGCCCGGTCTTAACCTGTTA	1680
Db	1774	GGTGCCCAACAACAAGAGGCGCGCGGGGAAAAAGGGGTGAAGGCCCGGTCTTAACCTGTTA	1833
QY	1681	CAA 1683 	
Db	1834	CAA 1836	
<hr/>			
RESULT 2			
AF093821                      2373 bp    mRNA       linear    ROD 04-OCT-1998			
LOCUS                      Mus musculus RRM RNA binding protein GRV-RBP mRNA, complete cds.			
DEFINITION                  AF093821			
VERSION                     AF093821.1 GI:3694985			
KEYWORDS			
SOURCE			
ORGANISM			
Mus musculus.			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
1 (bases 1 to 2373)			
Du G., Zhou Y., Chen J., Yuan J. and Qiang B.			
Direct Submission			
Submitted (23-SEP-1998) Department of Biochemistry and Molecular			
Biology, Institute of Basic Medical Sciences, 5 Dong Dan San Tiao,			
Beijing 100005, P.R. China			
Location/Qualifiers			
1..2373			
/organism="Mus musculus"			
/db_xref="taxon:10090"			
7..1884			
/codon_start=1			
/product="RRM RNA binding protein GRV-RBP"			
/protein_id="AAC62511.1"			
/db_xref="GI:3694986"			
/translation="MATEHAVNGNTEPEMDTTSAVIHSENFQTLLDAGLPKRVAKELD			
ETIYAVALHSDLDERRALBALKEFNEDDALAVLLOPKDSLSHYONKAFLCGWTKTY			
ROREKQGVTRVADSSGPDPEAKIKALEBTGYTLDTVOQRKYGPPSPSYVGSGOOPS			
GTEIVPGKLPRDLFDEBLVPEEKAPGLMDLRIMDDPLGLTLNRGYVAFTPTCKEAE			
AOEAALKNNHRIRSGKHIGVCIYANNRLTFGSIPEKTIKCOLLEESKYTEGLTVL			
ILTHNPDPDKKKRRCFPQIEDHKAAQAARRRKNSGKRYWNVGTVEKADPIEDPDE			
EMAVYKVLFRANLANVTBEITLESFSQFGKLERVKLKDYALTFBERGAYVAME			
EMNGDELGENETIEVFAPKPQKRKERARQAQNAQNYDYIYYGPHMPPTRGRR			
RGGRGAGVPDPYGEDYIDYIYGVDYHNRYGVEDPYGYEDFQVGARGGARAR			
AAPSRGAAPPGRGAGYSORGAPEGASGVARGAQOORGVARGARGGNVGGC			
RKADVENQPDTRRTQTNHMNSQPIAQQPLQGDHSNRYGKSBNQSFYDITGQQQ			
K"			
<hr/>			
BASE COUNT	740 A	412 C	587 G
ORIGIN			634 T
<hr/>			
Query Match                      95.2%; Score 1602; DB 10; Length 2373;			
Best Local Similarity 98.1%; Pred. No. 0;			
Matches 1646; Conservative 0; Mismatches 20; Indels 12; Gaps 2.			
QY	1	ATGGCTACGAACAATCTTAATGAAATGTACTGAAGACCACATGATACTACTTCAGCA	60
Db	7	ATGGCTACGAACAATCTTAATGAAATGTACTGAAGACCACATGATACTACTTCAGCA	66
QY	61	GTTATCATTCAGAAAATTTTTCAGATTCGCTTGAGCGCTTTACCAAGAAAGTTGCT	120
Db	67	GTTATCATTCAGAAAATTTTTCAGATTCGCTTGAGCGCTTTACCAAGAAAGTTGCT	126
QY	121	GA AAAA CTAGATGA AAATTTAC GTTGACAGGCGCTAG TTGCACATATGATTTA ATGTAAGA	180
Db	127	GA AAAA CTAGATGA AAATTTAC GTTGACAGGCGCTAG TTGCACATATGATTTA ATGTAAGA	186
QY	181	GCTATGAAGCTTTAA AAGATTCATGAAGAAGCGCGCAT TGCGAGTGCCTTCAACAGTTT	240

D	b	187	GCATCGAACGCTTTAAAAAGGCTCAATGAAAGCGCGCATTTGGCAGTGTCTTCAACAGCTT	246
O	y	241	AAAGA CAGTATCTTCTCA TGTTCAGAA CAAAGTGCCTTTTATATGTGAGTCATGAG	300
D	b	247	AAAGA CAGTATCTTCTCA TGTTCAGAA CAAAGTGCCTTTTATATGTGAGTCATGAG	306
O	y	301	ACTTACAGGCAGAGAGAAAAACAGGGGACCAAGTACAGACTCTAATGTAAGAACCTGAT	360
D	b	307	ACTTACAGGCAGAGAGAAAAACAGGGGACCAAGTACAGACTCTAATGTAAGAACCTGAT	366
O	y	361	GAGGCAAAAGATTAAAGGACTTTTGGAAAGAACAGGCTACACTGATGTGATCAATACAGT	420
D	b	367	GAGGCAAAAGATTAAAGGACTTTTGGAAAGAACAGGCTACACTGATGTGATCAATACAGT	426
O	y	421	CAGAAGAAATAGAGAGACCACTCCAGATCCGTTTATTCAGGTCAAGCAGCCTTCTGTT	480
D	b	427	CAGAAGAAATAGAGAGACCACTCCAGATCCGTTTATTCAGGTCAAGCAGCCTTCTGTT	486
O	y	481	GGCAGTGAATATTTGTGGGGAAAGATCCCAAGATCTGTTGAGGATGAGCTTGTTCGA	540
D	b	487	GGCAGTGAATATTTGTGGGGAAAGATCCCAAGATCTGTTGAGGATGAGCTTGTTCGA	546
O	y	541	TTATTTGAGAAAGCTGACCTATATAGGAACTTCGTTTAAATGAATGAGATCCG-----	591
D	b	547	TTATTTGAGAAAGCTGACCTATATAGGAACTTCGTTTAAATGAATGAGATCCGCTCACTGTT	606
O	y	592	CTCACTGGTCTCAACAGAGGTTATGCGTTTGTCACTTTTGTACAAAAGACGACACAA	651
D	b	607	CTCACTGGTCTCAACAGAGGTTATGCGTTTGTCACTTTTGTACAAAAGACGACCTCAG	666
O	y	652	GAGGCTGTTAACTGTATTAATATCATGAAATTCGTTCCGGAGACACATTTGGTGTGC	711
D	b	667	GAGGCTGTTAACTGTATTAATATCATGAAATTCGTTCCGGAGACACATTTGGTGTGC	726
O	y	712	ATCTCAGTGCCAACATATAGGCTTTTGTGGGCTGATTCCTAAGATGAACCAAGAG	771
D	b	727	ATCTCAGTGCCAACATATAGGCTTTTGTGGGCTGATTCCTAAGATGAACCAAGAG	786
O	y	772	CAGATTCTTGAGGAAATTATGCAAGTGA CAGAGGCTCA CAGATGTCA TTTATATACAC	831
D	b	787	CAGATTCTTGAGGAAATTATGCAAGTGA CAGAGGCTCA CAGATGTCA TTTATATACAC	846
O	y	832	CAACCTGATGACAAAGAAAAAAACAGAGGCTTTCTTGTGAATATGAAGATCACAAA	891
D	b	847	CAACCTGATGACAAAGAAAAAAACAGAGGCTTTCTTGTGAATATGAAGATCACAAA	906
O	y	892	ACAGCTCCCAAGGCACAGACGTAGGCTATATGATGCTAAAGTCAAAAGTCTGGGGAAATGTT	951
D	b	907	GCACCTCCCAAGGCACAGACGTAGGCTATATGATGCTAAAGTCAAAAGTCTGGGGAAATGTT	966
O	y	952	GAAACTGTGTAGTGGGGCTGATCCTATTGAAGATCCGATCCGAAAGTTATGGCAAAAGTGA	1011
D	b	967	GAAACTGTGTAGTGGGGCTGATCCTATTGAAGATCCGATCCGAAAGTTATGGCAAAAGTGA	1026
O	y	1012	AAAGTCTGTTTGTACGCAACTTTCGCAACAGGCTAACAGAGAAATTTTGAAGAACTCA	1071
D	b	1027	AAAGTCTGTTTGTACGCAACTTTCGCAACAGGCTAACAGAGAAATTTTGAAGAACTCA	1086
O	y	1072	TTTATGTCAGTTTGGGAACTCGAAACGATGAAGAACTTAAAGATATGCTTCATTCAT	1131
D	b	1087	TTTATGTCAGTTTGGGAACTCGAAACGATGAAGAACTTAAAGATATGCTTCATTCAT	1146
O	y	1132	TTTATGTCAGAGATGTGTGTGTGCAAGGCTATGGAAGAAATGATGTGAAGAACTTGAG	1191
D	b	1147	TTTATGTCAGAGATGTGTGTGTGCAAGGCTATGGAAGAAATGATGTGAAGAACTTGAG	1206
O	y	1192	GGAGAAATATTTGAATATGTTTTTGTCTAAGCCACAGATCTAGAAGAGGAAAGAAAGAAA	1251
D	b	1207	GGAGAAATATTTGAATATGTTTTTGTCTAAGCCACAGATCTAGAAGAGGAAAGAAAGAAA	1266
O	y	1252	GCTCAGAGGCAAGCAAGAAATCAATATATATATGATATCTACTATATATGTGTCACT	1311
D	b	1267	GCTCAGAGGCAAGCAAGAAATCAATATATATATGATATCTACTATATATGTGTCACT	1326

Oy	1312	CATATGCCCCCCCAACAAGAGGTGAGGGGCGGTGAGGTGAGGTGAGGTGAGGTGAGTATCCT	1371
Db	1327	CATATGCCCTCCCCCAACAAGAGGTGAGGGGCGGTGAGGTGAGGTGAGGTGAGTATCCT	1386
Oy	1372	CCAGATTATTTATGATACGAGATTATTTATGATTATTTATGATTATTAACCATTAAT	1431
Db	1387	CCAGATTATTTATGATACGAGATTATTTATGATTATTTATGATTATTAACCATTAAT	1446
Oy	1432	CGTGTGATATGAAGATCCATATCTATGTGTTATGAAGATTTTCAAGTTGAGCTAGAGA	1491
Db	1447	CGTGTGATATGAAGATCCATATCTATGTGTTATGAAGATTTTCAAGTTGAGCTAGAGA	1506
Oy	1492	AGGGGTGTGATAGAGAGCAAGAGGGGTGTCTCCATCCAGAGGTGTGTGGGCTCTCTCCC	1551
Db	1507	AGGGGTGTGATAGAGAGCAAGAGGGGTGTCTCTCCATCCAGAGGTGTGTGGGCTCTCTCCC	1566
Oy	1552	CGTGTGATAGAGCGGGTATTCACAGAGAGAGGCCCGTGATTCAGCAAGAGCGTTGGCGGT	1611
Db	1567	CGTGTGATAGAGCGGGTATTCACAGAGAGAGGCCCGTGATTCAGCAAGAGCGTTGGCGGT	1626
Oy	1612	GCGAGAGAGGTGCCCAACAACAAGAGGCCCGCGGGGAAAGAGGGGTGAGGCCGGT	1669
Db	1627	GCGAGAGAGGTGCC---CAACAAGAGGCCCGCGGGGTGAGGTGTCAGAGAGGGGTGCC	1681

RESULT 3				
AY034483				
LOCUS	AY034483	1666 bp	mRNA	linear
DEFINITION	Homo sapiens hnRNP Q1 mRNA, complete cds.			PRI 03-OCT-2001
ACCESSION	AY034483			
VERSION	AY034483.1 GI:15809589	-		

ORGANISM	Homo sapiens.
SOURCE	Homo sapiens.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1686)
TITLE	Mourrelatos,Z., Abdel,L., Yong,J., Kataoka,N. and Dreyfuss,G. SMN interacts with a novel family of hnRNP and spliceosomal proteins
JOURNAL	EMBO J. 20 (19), 5443-5452 (2001)
MEDLINE	21458434
PUBMED	11574476
REFERENCE	2 (bases 1 to 1686)
AUTHORS	Mourrelatos,Z.
TITLE	Direct Submission
JOURNAL	Submitted (11-MAY-2001) HMMI/Biochemistry, University of Pennsylvania, 330 CBH, 415 Curie Blvd., Philadelphia, PA 19104, USA
FEATURES	Location/Qualifiers
SOURCE	1. 1686

```

CDS
/Organism="Homo sapiens"
/db_xref="taxon:9606"
1.1686
/codon_start=1
/product="hnRN P Q1"
/protein_id="AAK5705.1"
/db_xref="gi:15809590"
/translation="MATEHWNGNGTEBMTDTSAVHSENPOTLLDAGLPQKVAEKD
EIVYAGVAHSDLDERALEALKEPNECDALVLPQPSDSLHQNKSAFLCYMKTY
RQRKQCTGVKSDSKSPDEAIKALLETGTTLVLTQQRKGPPDSVYSQQSPV
GTELFEVGIPEDELVEPLPEKQGPIDWRLIMDPLTGKTRAYALPCTEKAQOE
AVKLNNHIEIRSGHIGVCISVANNRFLVGSIPKSTQLLEBSKATBELTVILYI
HQPDDKKKRSFCPELEYEDHRTAAQARRSLSGVKTWNGGTEVEMADPLEDPEYTW
AKVVLFLRVNLANTVTEILEKAFSGFCKLERYKLDYAFIHFEEDGAVKAEENNG
KGLDEGENIEIVFPAKPPQKREKRAQCAALNMVDYDYYGGPHNPFRPGRGKNG
RGGQYPPEDYEDYDYYGYDYNVAGYGGEDPYVEDYEDVQARGGGAAGGAAP
SRGGAAPPRRAGSYDQGGGGRVARGVARGAGGAQQAQQRGRGVKEVAGPDLLQ"
BASE COUNT
536 a 282 c 435 g 433 t
Query Match 94.1%; Score 1583.8; DB 9; Length 1686;
Beet Locat Similarity 96.3%; Pred. No. 0;

```

Matches	1621;	Conservative	0;	Mismatches	62;	Indels	0;	Gaps	0;
Qy	1	ATGGCTACGAACATGTTTAATGAAATGCTGTAAGAGCCCATGATACTACTTCACGA	60						
Db	1	ATGGCTACGAACATGTTTAATGAAATGCTGTAAGAGCCCATGATACTACTTCGCA	60						
Qy	61	GTTATCCATTCAGAAAAATTTTTCAGACATGCGCTGATGCTGCTGTTTACACAGAAAGTGGCT	120						
Db	61	GTTATCCATTCAGAAAAATTTTTCAGACATGCTTGATGCTGGTTTACACAGAAAGTGGCT	120						
Qy	121	GA AAAA CTAGATGAATTTTACGTTGCA GGGCTAGTTGCA CAACTGATTTTATGATGAAGA	180						
Db	121	GA AAAA CTAGATGAATTTTACGTTGCA GGGCTAGTTGCA CAACTGATTTTATGATGAAGA	180						
Qy	181	GCTATCGAAGCTTTAAAGAGTTCAATGAAAGAGGGGCACTTGGCAGGCTTCAACAGTTT	240						
Db	181	GCTATCGAAGCTTTAAAGAGTTCAATGAAAGAGGGGCACTTGGCAGGCTTCAACAGTTT	240						
Qy	241	AAAGACAGTGCATCTCTCATGTTTCAAGAACAAAGTCCCTTTTATGTGAGTCAATAG	300						
Db	241	AAAGACAGTGCATCTCTCATGTTTCAAGAACAAAGTCCCTTTTATGTGAGTCAATAG	300						
Qy	301	ACTTACAGGCGAGAGAAAAACAGGGGACCAAGTACAGACTCTACTTAAGACCAAGAT	360						
Db	301	ACTTACAGGCGAGAGAAAAACAGGGGACCAAGTACAGACTCTACTTAAGACCAAGAT	360						
Qy	361	GAGGCAAAAGATTTAAGGCACTTTTGAAGAACAAGCTACACACTTGATGTGACTACAGT	420						
Db	361	GAGGCAAAAGATTTAAGGCACTTTTGAAGAACAAGCTACACACTTGATGTGACTACAGT	420						
Qy	421	CAGAGAAAGTATGAGAGACCACTCCAGATTCGCTTATTCAGTCA GACGCTTCTGTT	480						
Db	421	CAGAGAAAGTATGAGAGACCACTCCAGATTCGCTTATTCAGTCA GACGCTTCTGTT	480						
Qy	481	GGCACTGAGATTTTGTGGGAGAGATCCCAAGATGTGTTTGAAGATGAGCTTGTTC	540						
Db	481	GGCACTGAGATTTTGTGGGAGAGATCCCAAGATGTGTTTGAAGATGAGCTTGTTC	540						
Qy	541	TTATTTGAGAAAGCTGACCTATATGGAATCTTCGTTAA TGA TGAATCCGCTCACTG	600						
Db	541	TTATTTGAGAAAGCTGACCTATATGGAATCTTCGTTAA TGA TGAATCCGCTCACTG	600						
Qy	601	CTCAACAGAGTTATGCGGTTTGTCACTTTTGTGACAA AAGACAGACAAAGGCGT	660						
Db	601	CTCAATGAGGTTATGCGGTTTGTCACTTTTGTGACAA AAGACAGCTCAGAGGCGT	660						
Qy	661	AACTGTATATATATCATGAAATTCGTTCCGGGAGACACTTGGTGTCTCATCTCAGT	720						
Db	661	AACTGTATATATATCATGAAATTCGTTCCGGGAGACACTTGGTGTCTCATCTCAGT	720						
Qy	721	GCCAACATAGGCTTTTGTGGGCTCGATTCTTAAAGTAAACCAAGAGCAGATTC	780						
Db	721	GCCAACATAGGCTTTTGTGGGCTCGATTCTTAAAGTAAACCAAGAGCAGATTC	780						
Qy	781	GAGGAATTTAGCAAGTGA CAGAGGGTCTCA CAGATGTCATTTTATTCACCAACCTGAT	840						
Db	781	GAGGAATTTAGCAAGTGA CAGAGGGTCTTACAGAGCTCATTTTATTCACCAACCGGAT	840						
Qy	841	GACAAAGAAAAAACA GAGGCTTTTGTCTTCTTGAATATGAGATCA CAAA CAGCTGC	900						
Db	841	GACAAAGAAAAAACA GAGGCTTTTGTCTTCTTGAATATGAGATCA CAAA CAGCTGC	900						
Qy	901	CAGGCAAGATGATGCTTAATGATGTGTAAAGTCAAACTCTGGGGAAATGTTGGAATG	960						
Db	901	CAGGCAAGATGATGCTTAATGATGTGTAAAGTCAAACTCTGGGGAAATGTTGGAATG	960						
Qy	961	GAGTGGGCTGATCTTATTTGAAGATCTCGATCTCGAAGTTATGACCAAGTAAAGTGTCTG	1020						
Db	961	GAGTGGGCTGATCTTATTTGAAGATCTCGATCTCGAAGTTATGACCAAGTAAAGTGTCTG	1020						
Qy	1021	TTTGTACGCAACTTGTCCAA CAGGTTACAGAGAAATTTTGAAGAAAGTCATTTAAGTCA	1080						
Db	1021	TTTGTACGCAACTTGTCCAA TCTGTAA CAGAGAGATTTTGAAGAAAGTCATTTAAGTCA	1080						





	AY034481	AY034481	1872 bp	mRNA	linear	PRI 03-OCT-2001
LOCUS	AY034481	Homo sapiens hnRNP Q3 mRNA, complete cds.				
DEFINITION	AY034481					
ACCESSION	AY034481					
VERSION	AY034481.1	GI:15809585				
KEYWORDS						
SOURCE						
ORGANISM		Homo sapiens.				
		Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE		(bases 1 to 1872)				
AUTHORS		Mourales Z., Abel L., Yong J., Kataoka N. and Dreyfuss G.				
TITLE		SNM interacts with a novel family of hnRNP and spliceosomal proteins				
JOURNAL		EMBO J. 20 (19), 5443-5452 (2001)				
MEDLINE		21458434				
PubMed		11574476				
REFERENCE		2 (bases 1 to 1872)				
AUTHORS		Mourales Z.				
TITLE		Direct Submission				
JOURNAL		Submitted (11-MAY-2001) HMI/ Biochemistry, University of Pennsylvania, 330 CRB, 415 Curie Blvd., Philadelphia, PA 19104, USA				
FEATURES						
source		1..1872				
		/location=/qualifiers				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
CDS		1..1872				
		/codon_start=1				
		/product="hnRNP Q3"				
		/protein_id="AAK59703.1"				
		/db_xref="GI:15809586"				
		/translation="MATEHAVNGTGEPMDTSAVHSNFPQTLLDAGLFQKYAEKLD EIVYVAHSDLDERAIEALKEFNEDEGALAVIQFDSDLSHVONKSATLCGMVKTY ROSEQKTVDASSKGPEDAIKALLERTGYTLDTTGARKYGSPEDSYSGQPSV GTETPVCKIPIRLFEDELVPILFKAGIMPLRMLOPLTSLNGVAFTVCTEEABOE AVCLYNHHEIRSGKHICVCISVANNLRFVSITREKTKBQILLEPFSVTGLDVILY HQPDKKRSFCFLIEDHKTAQAQRRLMSKVWGVVEWADPLEDDPEVM AKVNFELVRNLANTVTEIELKAQSQGLMERVKQLKDVFHFDEMDGAVKMEEN KGLDEGENIEIVFAKPDPQKRERKARQCAAKOMYDYVYYPHPMPPTRRGGAG RGGVGPEDPYVEDYDYGVHYNRGRORGAAYOYGEDFOVGARRGGARGAGAP SRBGAGA PRPGRAGYSQRRGPGSARGRAGGAGAOORGGVARGAGRGAGRGKRX ADGNOPDSRKROTNNQMSQPIAQOPLLOGSHSGNYGKSENQERYDTFGQMK"				
BASE COUNT		550 a 323 c 489 g 470 t				
ORIGIN						
Query Match		92.2%; Score 1552.2; DB 9; Length 1872;				
Best Local Similarity		95.6%; Pred. No. Of				
Matches 1596; Conservative		0; Mismatches 73; Indels 0; Gaps 0;				
QY	1 ATGGCTACAGAACATGTTAAATGGAATGGTACTGAAGACCATTGACTTCTTCAGCA	60				
DB	1 ATGGCTACAGAACATGTTAAATGGAATGGTACTGAAGACCATTGACTTCTTCAGCA	60				
QY	61 GTTATCCATTCAAAAATTTTTCAGACATTTGCTTGATGCTGTTTACACAGAAAGTTGCT	120				
DB	61 GTTATCCATTCAAAAATTTTTCAGACATTTGCTTGATGCTGTTTACACAGAAAGTTGCT	120				
QY	121 GAAAAACATAGATTAATCGTTGCAGGGGCTAGTGTGACATAGATTTAGATGAAGA	180				
DB	121 GAAAAACATAGATTAATCGTTGCAGGGGCTAGTGTGACATAGATTTAGATGAAGA	180				
QY	181 GCTATCAGAGCTTTAAAGAGTTCAATGAACGCGCATTTGGCAGTGCCTTCAACAGTTT	240				
DB	181 GCTATCAGAGCTTTAAAGAGTTCAATGAACGCGCATTTGGCAGTGCCTTCAACAGTTT	240				
QY	241 AAAAGCAGTGAATCTCTCATGTTCAAGACAAAAGTGCTTTTTATGTGAGTCATGAAG	300				
DB	241 AAAAGCAGTGAATCTCTCATGTTCAAGACAAAAGTGCTTTTTATGTGAGTCATGAAG	300				
QY	301 ACTTACAGGACAGAGAAAAACAAGGACCAAAGTAGACAGACTTATGTAAGAAGCACAGAT	360				
DB	301 ACTTACAGGACAGAGAAAAACAAGGACCAAAGTAGACAGATTTATGTAAGAAGCACAGAT	360				



QY 361 GAGCAAAAGTAAAGCACTTTTGAAGAAAGAGCTACACATTGATGACTACAGT 420  
 DB 361 GAGGCAAAATTAAGCACTTTGGAAGAAAGAGCTACACATTGATGACTACAGT 420  
 QY 421 CAGAGGAATATGAGAGCACTTCAGATTCCTTTTATTCAGTCCAGAGCTTCTGTT 480  
 DB 421 CAGAGGAATATGAGAGCACTTCAGATTCCTTTTATTCAGTCCAGAGCTTCTGTT 480  
 QY 481 GGCATGAGATTTTGGGGGAGATCCCAAGATCTGTTTGAAGATGAGCTTGTCCA 540  
 DB 481 GGCATGAGATTTTGGGGGAGATCCCAAGATCTGTTTGAAGATGAGCTTGTCCA 540  
 QY 541 TTTATTTGAGAAAGCTGAGACCTATATGAGATCTTCTTTAATGATGATCCGCTCATCTGT 600  
 DB 541 TTTATTTGAGAAAGCTGAGACCTATATGAGATCTTCTTTAATGATGATCCGCTCATCTGT 600  
 QY 601 CTCAACAGAGGTTATGCGTTTGTCACTTTTGTACAAAGAGAGCAAGAGGCTGTT 660  
 DB 601 CTCAATATGAGGTTATGCGTTTGTCACTTTTGTACAAAGAGAGCTGAGAGGCTGTT 660  
 QY 661 AAATCTGATTAATATCATGAAATTCGTTCCGGGAGAGCAATGAGTCTGATCTCATGTT 720  
 DB 661 AAATCTGATTAATATCATGAAATTCGTTCCGGGAGAGCAATGAGTCTGATCTCATGTT 720  
 QY 721 GCCAACAATAGGCTTTTGTGGGCTCGATTCTTAAGATTAACCAAGAGAGCAATCTT 780  
 DB 721 GCCAACAATAGGCTTTTGTGGGCTCGATTCTTAAGATTAACCAAGAGAGCAATCTT 780  
 QY 781 GAGGAATTTAGCAAGAGTACAGAGGGCTCTCAAGATGTCATTTTATACCAACCACTGAT 840  
 DB 781 GAGGAATTTAGCAAGAGTACAGAGGGCTCTCAAGATGTCATTTTATACCAACCACTGAT 840  
 QY 841 GACAAGAAAAAAGAGAGCTTTTGTCTTCTTGAATATGAAATTCACAAACAGCTGCC 900  
 DB 841 GACAAGAAAAAAGAGAGCTTTTGTCTTCTTGAATATGAAATTCACAAACAGCTGCC 900  
 QY 901 CAGCAAGAGCTGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 DB 901 CAGCAAGAGCTGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 QY 961 GAGTGGGCTGATCTTATGAGATCTGATCTGAGATTTAAGCAAGAGTAAAGTGTCTG 1020  
 DB 961 GAGTGGGCTGATCTTATGAGATCTGATCTGAGATTTAAGCAAGAGTAAAGTGTCTG 1020  
 QY 1021 TTTGTAGCACTTGTGCAACAGGCTACAGAGAAATTTTGAAGATCATTTAGTCAAG 1080  
 DB 1021 TTTGTAGCACTTGTGCAACAGGCTACAGAGAAATTTTGAAGATCATTTAGTCAAG 1080  
 QY 1081 TTTGGGAACTGGAAGAGTAAAGAGCTAAAGATTAATGCTTCACTTTTGAAGAG 1140  
 DB 1081 TTTGGGAACTGGAAGAGTAAAGAGCTAAAGATTAATGCTTCACTTTTGAAGAG 1140  
 QY 1141 AGAGATGCTCTGCAAGGCTATGAGAAATGATGATGATGATGATGATGATGATGATGAT 1200  
 DB 1141 AGAGATGCTCTGCAAGGCTATGAGAAATGATGATGATGATGATGATGATGATGATGAT 1200  
 QY 1201 ATTGAATTTTGTGTAAGCCACAGATCAAGAGGAAAGAAAGAAAGAAAGAAAGAAAG 1260  
 DB 1201 ATTGAATTTTGTGTAAGCCACAGATCAAGAGGAAAGAAAGAAAGAAAGAAAGAAAG 1260  
 QY 1261 CAGCAGCAAGAGATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 DB 1261 CAGCAGCAAGAGATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 QY 1321 CCCCCAAGAGGCTGAGGCGGTGAGAGTGAAGTGGCTATGATGATGATGATGATGATGAT 1380  
 DB 1321 CCCCCAAGAGGCTGAGGCGGTGAGAGTGAAGTGGCTATGATGATGATGATGATGATGAT 1380  
 QY 1381 TATGATATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440  
 DB 1381 TATGATATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440

QY 1441 TATGATATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500  
 DB 1441 TATGATATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500  
 QY 1501 AGAGAGCAAGAGGCTCTGCTCCATCAGAGGCTGAGGAGCTCTCTCCCGTGTAGTA 1560  
 DB 1501 AGAGAGCAAGAGGCTCTGCTCCATCAGAGGCTGAGGAGCTCTCTCCCGTGTAGTA 1560  
 QY 1561 GCCGTTATTTACAGAGAGAGAGGCTCTGATCAGCAAGAGGCTTGCAGTCCAGAGGA 1620  
 DB 1561 GCCGTTATTTACAGAGAGAGAGGCTCTGATCAGCAAGAGGCTTGCAGTCCAGAGGA 1620  
 QY 1621 GGTGCCCAACAAGAGAGGCGCGGAGGAGAAAGAGGAGTGAAGGCGCGCTC 1669  
 DB 1621 GGTGCCCAACAAGAGAGGCGCGGAGGAGAAAGAGGAGTGAAGGCGCGCTC 1669  
 RESULT 6  
 AF037448  
 LOCUS  
 DEFINITION Homo sapiens RRM RNA binding protein GRY-rbp (GRY-RBP) mRNA,  
 complete cds.  
 ACCESSION AF037448  
 VERSION AF037448.1 GI:3037012  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE  
 1 (bases 1 to 2932)  
 Du G., Yuan J., Pan M., Yao H., Chen J. and Qiang B.  
 Cloning of a RRM RNA binding protein with glycine, arginine and  
 tyrosine-rich C-terminus  
 Unpublished  
 2 (bases 1 to 2932)  
 Du G., Yuan J., Pan M., Yao H., Chen J. and Qiang B.  
 Direct Submission  
 Submitted (09-DEC-1997) Department of Biochemistry and Molecular  
 Biology, Institute of Basic Medical Sciences, 5 Dong Dan San Tiao,  
 Beijing 100005, P.R. China  
 JOURNAL  
 TITLE  
 AUTHORS  
 Du G., Yuan J., Pan M., Yao H., Chen J. and Qiang B.  
 DIRECT SUBMISSION  
 JOURNAL  
 TITLE  
 AUTHORS  
 Du G., Yuan J., Pan M., Yao H., Chen J. and Qiang B.  
 FEATURES  
 source  
 1..2932  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 1..2932  
 /gene="GRY-RBP"  
 526..2397  
 /gene="GRY-RBP"  
 /note="RRM RNA binding protein; RNA recognition motif  
 binding protein; glycine, arginine and tyrosine-rich at  
 C-terminus"  
 /codon\_start=1  
 /product="GRY-rbp"  
 /protein\_id="AAC12926.1"  
 /db\_xref="GI:3037013"  
 /translation="MATEHVNGGEBEPMDTTSVISHENRQTLIDAGLPQVAKLD  
 ELYAGLVASHDLDERALBALREPEDBALVLOQPKOSDLSHVONKSAFLGVKTY  
 ROREKQKRVADSSKGPDEBAIKALERTGYTLDTGQKRGPPSPVSGQPSV  
 GTEIFVGKIPDLFEDBLVLEKAGPIWDLRLMDPLGLNRGYAFVPTKEAAOE  
 AVKLNNHISGSKHIGVICIVANNRLFVGSIPKSKTEQLIEBSKYTEGLDIVILY  
 HOPDKKNSRSCPLEYEDHRTAQAARRLSGKVKWNGVGVEMADPIEDPEVM  
 AKVYLFTPLNANTYTERILBKASQPKLERVKLKYQYAFHPBERGAYKAMEEM  
 KGDLEGENIEIVFAKPPDQKREKRAQQAQANQMTDYIYYPHPMPPTKRGKG  
 RGQVYPPDYGYEDYDYDYGYDYVYNYGAGYEDPYGYEDYQVAGRGGRARQAP  
 SRGRGAPRGAGYSORSGPSAGVGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  
 ADGVNQPDSKROTNNQWMSOPIAQCPLAGGDHSGNGYKSENQEFYQDFGQGWK"  
 BASE COUNT 868 a 520 c 823 g 721 t  
 ORIGIN  
 Query Match 92.2%; Score 1552.2; DB 9; Length 2932;  
 Best Local Similarity 95.6%; Pred. No. 0;  
 Matches 1596; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 1 ATGGCTACAGAACATGTTAATGSAATGGTACTGAGAGCCCATGATCTACTTTCAGCA 60  
 DB 526 ATGGCTACAGAACATGTTAATGSAATGGTACTGAGAGCCCATGATCTACTTTCAGCA 585  
 QY 61 GTTATCCATTGAGAAATTTTTCAGACATTTGCTGTTGTTTACCAAGAAAGTTGCT 120  
 DB 586 GTTATCCATTGAGAAATTTTTCAGACATTTGCTGTTGTTTACCAAGAAAGTTGCT 645  
 QY 121 GAAAACTGATGAAATTTTTCAGGAGGCTAGTTCACATGATGTTTGAAGAA 180  
 DB 646 GAAAACTGATGAAATTTTTCAGGAGGCTAGTTCACATGATGTTTGAAGAA 705  
 QY 181 GCTATCGAAGCTTTTAAAGAGTTCAATGAGCGGCGATTTGGCAGTCTTCAAGTTT 240  
 DB 706 GCTATCGAAGCTTTTAAAGAGTTCAATGAGCGGCGATTTGGCAGTCTTCAAGTTT 765  
 QY 241 AAGACAGTGTCTCTCATGTTTCAAGACAAAGTGCCTTTTATGAGATCATGAG 300  
 DB 766 AAGACAGTGTCTCTCATGTTTCAAGACAAAGTGCCTTTTATGAGATCATGAG 825  
 QY 301 ACTTACAGGAGAGAGAAACAGGGGACCAAGTACAGACTCTGATTAAGACAGAT 360  
 DB 826 ACTTACAGGAGAGAGAAACAGGGGACCAAGTACAGACTCTGATTAAGACAGAT 885  
 QY 361 GAGGCAAAAGATTAAAGGCACTTTTGGAAAGAAAGGCTACACACTGATGTAGTCAAGGT 420  
 DB 886 GAGGCAAAAGATTAAAGGCACTTTTGGAAAGAAAGGCTACACACTGATGTAGTCAAGGT 945  
 QY 421 CAGAGAAATTTGAGAGACACCTTCAGATTCGTTTATTCAGGTCAAGCAGCTTCTGTT 480  
 DB 946 CAGAGAAATTTGAGAGACACCTTCAGATTCGTTTATTCAGGTCAAGCAGCTTCTGTT 1005  
 QY 481 GGCCTGAGATTTTGGGGGAAATCCCGAAGATCTGTTTGAAGATGAGCTTGTCCA 540  
 DB 1006 GGCCTGAGATTTTGGGGGAAATCCCGAAGATCTGTTTGAAGATGAGCTTGTCCA 1065  
 QY 541 TTATTTGAGAAAGCTGACCTATATGAGGATCTTCTTATATGATGATCCGCTACAGGT 600  
 DB 1066 TTATTTGAGAAAGCTGACCTATATGAGGATCTTCTTATATGATGATCCGCTACAGGT 1125  
 QY 601 CTCAACAGAGGTTATGCGTTTGTCACTTTTGTCAAAAAGAGCAGCAAGAGGCTGTT 660  
 DB 1126 CTCAACAGAGGTTATGCGTTTGTCACTTTTGTCAAAAAGAGCAGCTCAAGAGGCTGTT 1185  
 QY 661 AAACGTATATATATCAAGAAATTCCTCCGGAAGCAGATTTGGTGTGATCTCAGTT 720  
 DB 1186 AAACGTATATATATCAAGAAATTCCTCCGGAAGCAGATTTGGTGTGATCTCAGTT 1245  
 QY 721 GCCAACAATAGGCTTTTGTGGGCTCGATTCCTAAGATTAACCAAGAGCAGATTCCT 780  
 DB 1246 GCCAACAATAGGCTTTTGTGGGCTCGATTCCTAAGATTAACCAAGAGCAGATTCCT 1305  
 QY 781 GAGGAATTTAGCAAAAGTGACAGAGGCTCAAGATGTCATTTTATACCAACCAACTGAT 840  
 DB 1306 GAGGAATTTAGCAAAAGTGACAGAGGCTCAAGATGTCATTTTATACCAACCAACTGAT 1365  
 QY 841 GACAAAGAAAAAAGAGGCTTTTGTCTTCTTGAATTAAGATCAAAAAGCTGCC 900  
 DB 1366 GACAAAGAAAAAAGAGGCTTTTGTCTTCTTGAATTAAGATCAAAAAGCTGCC 1425  
 QY 901 CAGCAAGACGTAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 DB 1426 CAGCAAGACGTAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1485  
 QY 961 GAGTGGGCTGATCTTATTTAAAGATCTGATCTGAAAGTTATGCGCAAAAGTAAAGTCTG 1020  
 DB 1486 GAGTGGGCTGATCTTATTTAAAGATCTGATCTGAAAGTTATGCGCAAAAGTAAAGTCTG 1545  
 QY 1021 TTGTATCCCACTTTGCTCAACAGGTAAACAGAAATTTTGAAGAAAGCTTTATGTCAG 1080  
 DB 1546 TTGTATCCCACTTTGCTCAACAGGTAAACAGAAATTTTGAAGAAAGCTTTATGTCAG 1605  
 QY 1081 TTGTGGAAACTGGAACGATTAAGAAAGTAAAGATTAAGTTCATTTTATGATGAG 1140

DB 1606 TTGTGGAAACTGGAACAGATGGAAGAGTTAAAGATTAAGCTTCAATTCATTTGATGAG 1665  
 QY 1141 AGAATGCTGCTGTCAAGGCTATGGAAGAAATGATGATTAAGACTTGGAGGAGAAAT 1200  
 DB 1666 CGAATGCTGCTGTCAAGGCTATGGAAGAAATGATGATTAAGACTTGGAGGAGAAAT 1725  
 QY 1201 ATTGAATTTGTTTGTAAAGCCACAGATCAGAAAGAAAGAAAGAAAGCTCAGAG 1260  
 DB 1726 ATTGAATTTGTTTGTCAAGCCACAGATCAGAAAGAAAGAAAGAAAGCTCAGAG 1785  
 QY 1261 CAAGCAGCAAAAGATCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 DB 1786 CAAGCAGCAAAAGATCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1845  
 QY 1321 CCCCACAAAGAGTGAAGGCGTGAAGGTGAGAGGTGCTATGATATCTCCAGATTAAT 1380  
 DB 1846 CTTCCAAAGAGAGTGAAGGCGTGAAGGTGAGAGGTGCTATGATATCTCCAGATTAAT 1905  
 QY 1381 TATGATACGAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1440  
 DB 1906 TATGATACGAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1965  
 QY 1441 TATGAAGATCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
 DB 1966 TATGAAGATCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2025  
 QY 1501 AGAGAGCAAGGAGTGTGCTCATCAGAGGTCTGAGGCTGCTCTCCCTGTGTAGA 1560  
 DB 2026 AGAGAGCAAGGAGTGTGCTCATCAGAGGTCTGAGGCTGCTCTCCCTGTGTAGA 2085  
 QY 1561 GCGGTTATTTCAAGAGAGAGGCTCTGATCAGCAAGAGGCTTGGCGTGTGAAGGA 1620  
 DB 2086 GCGGTTATTTCAAGAGAGAGGCTCTGATCAGCAAGAGGCTTGGCGTGTGAAGGA 2145  
 QY 1621 GGTGCCCAACAAAGAGGCGCGGGGAGAAAGGGGTGAGGCGGCTC 1669  
 DB 2146 GGTGCCCAACAAAGAGGCGCGGGGAGAAAGGGGTGAGGCGGCTC 2194

RESULT 7  
 AR177851  
 LOCUS AR177851 2079 bp DNA linear PART 17-DEC-2001  
 DEFINITION Sequence 2 from patent US 6313266.  
 ACCESSION AR177851  
 VERSION AR177851.1 GI:17920206  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 2079)  
 AUTHORS Bandman, O., Yue, H., Corley, N.C. and Shah, P.  
 TITLES Human nucleolin-like protein  
 JOURNAL Patent: US 6313266-A 2 06-NOV-2001;  
 FEATURES  
 source 1..2079  
 BASE COUNT 639 a 377 c 535 g 528 t  
 ORIGIN  
 Query Match 90.3%; Score 1519; DB 6; Length 2079;  
 Best Local Similarity 96.0%; Pred. No. 0;  
 Matches 1558; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
 QY 1 ATGGCTACAGAACATGTTAATGSAATGGTACTGAGAGCCCATGATCTACTTTCAGCA 60  
 DB 203 ATGGCTACAGAACATGTTAATGSAATGGTACTGAGAGCCCATGATCTACTTTCAGCA 262  
 QY 61 GTTATCCATTGAGAAATTTTTCAGACATTTGCTGTTTACCAAGAAAGTTGCT 120  
 DB 263 GTTATCCATTGAGAAATTTTTCAGACATTTGCTGTTTACCAAGAAAGTTGCT 322  
 QY 121 GAAAACTGATGAAATTTTTCAGGAGGCTAGTTCACATGATGTTTGAAGAA 180

```

Db 323 GAAAACTAGATGAAATTTGCTGAGGCGTGTGACATAGTGAATTAAGAAAGA 382
Qy 181 GCTATGAGACCTTAAAGAGTTCAATGAAGCGCGCATTTGGAGTCTTCAACAGTT 240
Db 383 GCTATGAGACCTTAAAGAGTTCAATGAAGCGCGCATTTGGAGTCTTCAACAGTT 442
Qy 241 AAGACAGTATCTCTCATGTTTCAAGAACAAAGTCTTTTATATGAGATCATAG 300
Db 443 AAGACAGTATCTCTCATGTTTCAAGAACAAAGTCTTTTATATGAGATCATAG 502
Qy 301 ACTTACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 503 ACTTACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 562
Qy 361 GAGGCAAGATTAAAGCACTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 563 GAGGCAAGATTAAAGCACTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 622
Qy 421 CAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 623 CAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 682
Qy 481 GGCATGAGATATTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 683 GGCATGAGATATTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 742
Qy 541 TATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 743 TATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 802
Qy 601 CTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 803 CTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 862
Qy 661 AAATCTATATATATATATATATATATATATATATATATATATATATAT 720
Db 863 AAATCTATATATATATATATATATATATATATATATATATATATATAT 922
Qy 721 GCCAACAATAGGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 923 GCCAACAATAGGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 982
Qy 781 GAGGAATTTAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 983 GAGGAATTTAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1042
Qy 841 GACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 1043 GACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1102
Qy 901 CAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 1103 CAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1162
Qy 961 GAGTGGGCTGATCTCTATTTGAAGATCTCTGAGAGAGAGAGAGAGAGAGAG 1020
Db 1163 GAGTGGGCTGATCTCTATTTGAAGATCTCTGAGAGAGAGAGAGAGAGAGAG 1222
Qy 1021 TTTTGAAGCACTTTGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1223 TTTTGAAGCACTTTGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1282
Qy 1081 TTTTGAAGCACTTTGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1283 TTTTGAAGCACTTTGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1342
Qy 1141 AGAGATGAGTCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1343 CGAGATGAGTCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1402
Qy 1201 ATTGAATTTGTTTGTGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260

```

```

Db 1403 ATTGAATTTGTTTGTGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1462
Qy 1261 CAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1463 CAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1522
Qy 1321 CCCCACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1523 CCCCACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1582
Qy 1381 TATGATTAAGAGATTTATGATTTATGATTTATGATTTATGATTTATGATTTAT 1440
Db 1583 TATGATTAAGAGATTTATGATTTATGATTTATGATTTATGATTTATGATTTAT 1642
Qy 1441 TATGATTAAGAGATTTATGATTTATGATTTATGATTTATGATTTATGATTTAT 1500
Db 1643 TATGATTAAGAGATTTATGATTTATGATTTATGATTTATGATTTATGATTTAT 1702
Qy 1501 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 1703 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1762
Qy 1561 GCCGTTATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1763 GCCGTTATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1822
Qy 1621 GGT 1623
Db 1823 GGT 1825

RESULT 8
AF408434 2232 bp mRNA linear ROD 10-OCT-2001
AF408434
DEFINITION Mus musculus RRM RNA binding protein NSAP1 (Nsap1) mRNA, partial cds.
ACCESSION AF408434
VERSION AF408434.1 GI:15986734
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 2232)
AUTHORS Zhou, M. and Raschke, W.C.
TITLE Mus musculus RRM RNA binding protein, NSAP1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2232)
AUTHORS Zhou, M. and Raschke, W.C.
TITLE Direct Submission
JOURNAL Submitted (13-AUG-2001) Immunology, Sidney Kimmel Cancer Center, 10835 Altman Row, San Diego, CA 92121, USA
FEATURES
source
1..2232
/organism="Mus musculus"
/db_xref="taxon:10090"
<1..2232
/gene="Nsap1"
<1..1679
/gene="Nsap1"
/codon_start=3
/product="RRM RNA binding protein NSAP1"
/protein_id="AA11726.1"
/db_xref="GI:15986734"
translation="KEPNDGALAVLQDFKDSLSHYNKSAPLGGWKTYPORPKOG
TKVADSQGPDEAKTKALLERTGTLDTVQGRKGPSPPSVYSQQPSVGTETPVG
KIPRLPDELVLPEKAGPTMDRLMMDPLTGLNRGTFVFTCKEAAQAVKLYNN
HIRSGKHIVGICISVANNRLRVGSIPIKSTKEQLIEBSKYTBGLTDVILVHOPDKK
NRGFCLEVEDHHTAAQARRRLMSGKTVGNGVTVEMADPIEDPDEVAKYVLF
VRNLANTVEIEILKSFQFGLSRVKKLKOYAFIHPERGAAYVAMEENKGGKLEGE
NIEIVFAKPEPKRERKRAORQAAKNQYDDYVYGGPHMPPTRGRRGGAGGYGP
PDYGYENYVYVGYDYNRYGVEDPYGYGEDPYONCARGGGAGARAAASRRGAA
PPRGAGISGNGPESANGVARGAGAAQDQGRGVARGRGAGVGGKRYADGNP

```



## FEATURES

source

http://www.sanger.ac.uk/Projects/C\_elegans/wormp/ This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20> RP4-775C13 is from the library RP4-4 constructed by the group of Pletier de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCVPAC2

This sequence is the entire insert of clone RP4-775C13.

Location/Qualifiers

```
1..103819
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="p12.1-13"
/clone="RP4-775C13"
/clone_lib="RP4-4"
1..236
/note="AluY repeat: matches 1..236 of consensus"
repeat_region
293..562
/note="L1M4c repeat: matches 1844..1624 of consensus"
repeat_region
819..1471
/note="L1M3c repeat: matches 969..1634 of consensus"
repeat_region
2177..2588
/note="L1M3c repeat: matches 14..431 of consensus"
repeat_region
2593..2652
/note="MUTIF repeat: matches 474..533 of consensus"
repeat_region
2701..3176
/note="238 copies 2 mer tt 54% conserved"
repeat_region
2814..2978
/note="15 copies 11 mer 66% conserved"
repeat_region
2996..3196
/note="67 copies 3 mer tcc 56% conserved"
repeat_region
3031..3184
/note="14 copies 11 mer 63% conserved"
repeat_region
3296..4082
/note="L1M4 repeat: matches 3350..4128 of consensus"
repeat_region
4197..4315
/note="L1M4 repeat: matches 3021..3136 of consensus"
repeat_region
4316..4738
/note="MSTA repeat: matches 1..426 of consensus"
repeat_region
4739..5118
/note="L1M4 repeat: matches 2626..3021 of consensus"
repeat_region
5146..5288
/note="match: GSS: Em:B90786"
misc_feature
5345..6619
/note="L1PA16 repeat: matches 3863..5135 of consensus"
repeat_region
6882..7139
/note="129 copies 2 mer gg 58% conserved"
repeat_region
7351..8891
/gene="dJ775C13.1"
7351..8891
/gene="dJ775C13.1"
/gene="dJ775C13.1"
ribonuclease R, RRM rna binding protein GRX-RBP)
pseudogene)
match: CDNA5: Em:AF000364 Em:AF037448 Em:AF03821
match: ESTs: Em:AA41147 Em:W38408 Em:AI765128 Em:AV147169
Em:AA411148 Em:AA307374 Em:T69820 Em:AA677426 Em:T89909
Em:AI682619 Em:AI472757 Em:AI548569 Em:AV134944
Em:AA912195 Em:AA094744 Em:AI084864 Em:H34275 Em:AA03140
Em:AI289939 Em:ID76738 Em:AA16046 Em:T39242 Em:AA407580
Em:AA561471 Em:AA647246 Em:AI115211 Em:AA565439
Em:AI581394 Em:H32663 Em:AI566089 Em:AA207215 Em:AA216046
Em:T35450 Em:AA37033 Em:AA465413 Em:W29501 Em:M61969
Em:AA164877 Em:AI625087 Em:AA465343 Em:AA524815
Em:AA563205 Em:AA066208 Em:AA045158 Em:AA524815
Em:AA544174 Em:AA18195 Em:AA664286 Em:AI64876
Em:AI267372 Em:AA150302 Em:AA146001 Em:AA662366
Em:AA130625 Em:AA222633 Em:AA242157 Em:AA130793
Em:AA654252 Em:AA306863 Em:AA186327 Em:AA153225
```

```
misc_feature
7549..8004
/gene="dJ775C13.1"
/note="match: GSS: Em:AQ434110"
7614..7732
/gene="dJ775C13.1"
/note="match: GSS: Em:AQ536515"
complement(8069..8562)
/note="match: STS: Em:T89722"
complement(8476..9063)
/note="match: GSS: Em:AQ008463"
8721..8816
/note="32 copies 3 mer att 65% conserved"
repeat_region
9341..10347
/note="L1PA16 repeat: matches 5128..6150 of consensus"
repeat_region
10381..10517
/note="L1M4 repeat: matches 2211..2348 of consensus"
repeat_region
11105..11201
/note="HAL1 repeat: matches 675..775 of consensus"
11261..11455
/note="L2 repeat: matches 645..869 of consensus"
11456..11765
/note="L1M4 repeat: matches 1..307 of consensus"
11766..11863
/note="L2 repeat: matches 555..645 of consensus"
12356..12519
/note="L2 repeat: matches 298..464 of consensus"
12520..12615
/note="L1M4 repeat: matches 6051..6146 of consensus"
13257..13396
/note="L1M4 repeat: matches 6051..6146 of consensus"
13442..13893
/note="MERS4B repeat: matches 1..498 of consensus"
13894..13931
/note="19 copies 2 mer ac 89% conserved"
repeat_region
14475..14910
/note="L1M4 repeat: matches 5707..6148 of consensus"
14960..15480
/note="L1M4 repeat: matches 4030..4551 of consensus"
15496..16282
/note="L1PA4 repeat: matches 5361..6146 of consensus"
16286..16427
/note="L1PA4 repeat: matches 5213..5558 of consensus"
16428..16958
/note="match: GSS: Em:AQ427784"
17470..18240
/note="L1PB1 repeat: matches 5277..6151 of consensus"
18383..18706
/note="AluY repeat: matches 1..309 of consensus"
18793..18837
/note="15 copies 3 mer att 75% conserved"
repeat_region
19213..19509
/note="AluY repeat: matches 3..298 of consensus"
complement(join(19771..19905,51843..51912))
/note="match: GSS: Em:AQ377847"
22426..22613
/note="MERS3 repeat: matches 4..188 of consensus"
```

```

repeat_region 2385..23561
/note="MER5B repeat: matches 1..178 of consensus"
repeat_region 23821..23871
/note="L2 repeat: matches 245..2496 of consensus"
repeat_region 23899..24082
/note="MER5A repeat: matches 1..181 of consensus"
repeat_region 24363..24676
/note="AluX repeat: matches 3..312 of consensus"
repeat_region 25302..25481
/note="MLTIG repeat: matches 9..187 of consensus"
repeat_region 25486..25537
/note="26 copies 2 mer ct 86% conserved"
repeat_region 25564..25591
/note="MLTIG repeat: matches 243..270 of consensus"
repeat_region 26477..26762
/note="143 copies 2 mer tc 56% conserved"
repeat_region 26510..26718
/note="19 copies 11 mer 66% conserved"
repeat_region 26836..27041

Query Match 85.7%; Score 1442.8; DB 9; Length 103819;
Best Local Similarity 94.1%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 87; Indels 13; Gaps 8;

Qy 1 ATGGCTACAGAACATGTATAGAAATGTACTGGAAGAGCCCATGAGATCTACTTACAGCA 60
Db 7351 ATGGCTACAGAACATGTATAGAAATGTACTGGAAGAGCCCATGAGATCTACTTACAGCA 7410
Qy 61 GTTATCCATTAGAAATTTTTCAGCATTTGCTGTGATGCTGTTTACACAGAAAGTTGCT 120
Db 7411 GTTATCCATTAGAAATTTTTCAGCATTTGCTGTGATGCTGTTTACACAGAAAGTTGCT 7470
Qy 121 GAAACATAGTAAATTTAGTTCAGAGGCTAGTGCACATAGATGATTTAGTAAAGA 180
Db 7471 GAAACATAGTAAATTTAGTTCAGAGGCTAGTGCACATAGATGATTTAGTAAAGA 7530
Qy 181 GCTATCGAAGCTTTAAAGAGTTCAATAGAGCGGCGATGGGAGTCTTCAACAGTTT 240
Db 7531 GCTATCGAAGCTTTAAAGAGTTCAATAGAGCGGCGATGGGAGTCTTCAACAG-TT 7589
Qy 241 AAGACAGTGTCTCTCATGTTCAGAACAAAGTGCCTTTTATGTGAGTCATGAG 300
Db 7590 AAGACAGTGTCTCTCGTTCAGAACAAAGTGCCTTTTATGTGAGTCATGAGAA 7649
Qy 301 ACTTACGCGGAGAGAGAAACAGGGGACCAAGTAGAGACTCTAGTAAAGACCAAT 360
Db 7650 CTTATAGGCGAGAGAGAAACAGGGGACCAAGTAGAGACTCTAGTAAAGACCAAT 7709
Qy 361 GAGGCAAGATTTAAGGCACTTTTGAAGAAACAGGCTACACATTTGATGTAGTACAGT 420
Db 7710 GAGGCAAGATTTAAGGCACTTTTGAAGAAACAGGCTACACATTTGATGTAGTACAGT 7769
Qy 421 CAGAGAGATAGAGACCACTTCAGATTCGTTTATTCAGGTCAGACAGCTTCTGTT 480
Db 7770 CAGAGAGATAGAGACCACTTCAGATTCGTTTATTCAGGTCAGACAGCTTCTGTT 7829
Qy 481 GGCATGAGATTTTGTGGGGAAGATCCCCAGAGATCTGTTTGAAGATGAGCTTGTCCA 540
Db 7830 GGCATGAGATTTTGTGGGGAAGATCCCCAGAGATCTGTTTGAAGATGAGCTTGTCCA 7889
Qy 541 TTATTTGAGAAGCTGACCTATATAGGATCTTGTTTATGATGATCGGTACAGTGT 600
Db 7890 TTATTTGAGAAGCTGACCTATATAGGATCTTGTTTATGATGATCGGTACAGTGT 7949
Qy 601 CTCAACAGAGTTATGCTTTGTCACTTTTGTACAAAGAGACAGCAAGAGGCTGTT 660
Db 7950 CTCAACAGAGTTATGCTTTGTCACTTTTGTACAAAGAGAGCTCGGAGAGCTG-TTA 8007
Qy 661 AAACGTATATATCATGAATTTGTCGCGAAGACATTTGCTGCTCACTTCACTT 720
Db 8008 AAACGTATATATCATGAATTTGTCGCGAAGACATTTGCTGCTCACTTCACTT 8067
Qy 721 GCCAACATAGGCTTTTGTGGGCTGATCTCTAAGAG-TAAACCAAGAGACAGATCTT 779

```

```

Db 8068 GCCAACATAGGCTTTTGTGGGCTGATCTCTAAGAGTTAAACCAAGAGACAGATCTT 8127
Qy 780 TGAAGATTTAGC-AAAGTAGAGAGGCTCTACAGATGTCATTTATACCAACCTTG 838
Db 8128 TGAAGATTTAGCAAAAGTAAAGAGAGGCTCTACAGATGTCATTTATACCAACCTTG 8187
Qy 839 ATGACAAAGAAAAAACAAGAGGCTTTTGTCTTTCTTAATATGAAGATCCAAACAGCTG 898
Db 8188 ATGACAAAGAAAAAACA-AGGCTTTTGTCTTTCTTAATATGAAGATCCAAACAGCTG 8245
Qy 899 CCGAGCAAGCTAGGCTATAGTGTGTAAGTCAAGTCTGGGGAATGTTGAACTG 958
Db 8246 CCGAGCAAGCTAGGCTATAGTGTGTAAGTCAAGTCTGGGGAATGTTGAACTG 8305
Qy 959 TTAGAGGCTGATCTTATTAAGATCTGATCTGAAATGATAGCAAGATTAAGAGTGC 1018
Db 8306 TTAGAGGCTGATCTTATTAAGATCTGATCTGAAATGATAGCAAGATTAAGAGTGC 8365
Qy 1019 TGTTTGACCAACCTTGCCACACGGTAAACAGAGAAATTTTGAAGATCTTATAGTC 1078
Db 8366 TGTTTGACCAACCTTGCCACACGGTAAACAGAGAAATTTTGAAGATCTTATAGTC 8425
Qy 1079 AGTTGGAAACTGGAACGAGTAAAGGCTAAAGATTAATGCTTTCATTCATTTGATG 1138
Db 8426 AGTTGGAAACTGGAACGAGTAAAGGCTAAAGATTAATGCTTTCATTCATTTGATG 8485
Qy 1139 AGAGAGATGCTGCTGTCAGGCTATAGAGAAATGAATGTAAAGACTTGAGAGGAGAA 1198
Db 8486 AGAGAGATGCTGCTGTCAGGCTATAGAGAAATGAATGTAAAGACTTGAGAGGAGAA 8545
Qy 1199 ATATGAAATTTTGTTCCTAAGCCACAGATCAGAGAGAGAAAGAAAGCTCAGA 1258
Db 8546 ATATGAAATTTTGTTCCTAAGCCACAGATCAGAGAGAGAAAGAAAGAA-AAAAGCTCAGA 8603
Qy 1259 GGCAGAGCAAGAAATCAATGATGATGATTAATTAATTAATGCTCAGCTCATATGC 1318
Db 8604 GGCAGAGCAAGAAATCAATGATGATGATTAATTAATTAATGCTCAGCTCATATGC 8663
Qy 1319 CTCCTCCCAAGAGAGTGAAGGCGTGAAGTGAAGGCTGATGATATCTCCAGATT 1378
Db 8664 CTCCTCCCAAGAGAGTGAAGGCGTGAAGTGAAGGCTGATGATATCTCCAGATT 8723
Qy 1379 ATATGATACGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1438
Db 8724 ATATGATACGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8783
Qy 1439 GATATGAATCACTATGATGATTAATGAAGATTTCAAGTTGAGCTAGAGAGAGGCTG 1498
Db 8784 GATATGAATCACTATGATGATTAATGAAGATTTCAAGTTGAGCTAGAGAGAGGCTG 8842
Qy 1499 GTAGAGAGCAAGAGGCTGCTCATTCAGAGGCTGAGGAGGCTGCTCCCGTGGTA 1558
Db 8843 GTAGAGAGCAAGAGGCTGCTCATTCAGAGGCTGAGGAGGCTGCTCCCGTGGTA 8902
Qy 1559 GAGCGGCTTATTCACAGAGAGAGAGGCTTGAATCAGCAAGAGGCTTCCGAGTCCAGAG 1618
Db 8903 GAGCGGCTTATTCACAGAGAGAGAGGCTTGAATCAGCAAGAGGCTTCCGAGTCCAGAG 8962
Qy 1619 GAGGTGCCCAACAAAGAGAGGCGCGG---GGGAAAAAGGGGTGAGAGCGGCTCGAGC 1675
Db 8963 GAGGTGCCCAACAAAGAGAGGCGCGGCGGAGGAGGAGGAGGCTGAGGCTCGAGC 9022
Qy 1676 TGTTCACA 1683
Db 9023 TGTTCACA 9030

RESULT 10
AL390737 212659 bp DNA linear HTG 21-MAR-2001
LOCUS Homo sapiens chromosome 13 clone RP11-190P17, *** SEQUENCING IN
DEFINITION PROGRESS ***; 16 unordered pieces.

```

ACCESSION AL390737  
VERSION AL390737.9 GI:13398796  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 212659)  
REFERENCE  
AUTHORS Burton, J.  
TITLE Direct Submission  
JOURNAL Submitted (17-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk  
On Mar 21, 2001 this sequence version replaced gi:13396612.  
----- Genome Center  
Center: Sanger Centre  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquerry@sanger.ac.uk  
----- Project Information  
Center project name: BA190P17  
----- Summary Statistics  
Assembly program: XGAP4, version 4.5  
Sequencing vector: plasmid; L08752; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Consensus quality: 203924 bases at least Q40  
Consensus quality: 207400 bases at least Q30  
Consensus quality: 209306 bases at least Q20  
Insert size: 21159; sum-of-contigs  
Quality coverage: 7.54x in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 16 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 48415: contig of 48415 bp in length  
48416 48515: gap of 100 bp  
48516 82643: contig of 34128 bp in length  
82644 82743: gap of 100 bp  
82744 89101: contig of 6358 bp in length  
89102 89201: gap of 100 bp  
89202 127367: contig of 38166 bp in length  
127368 127467: gap of 100 bp  
127468 169067: contig of 41600 bp in length  
169068 169167: gap of 100 bp  
169168 171259: contig of 2092 bp in length  
171260 171359: gap of 100 bp  
171360 173612: contig of 2253 bp in length  
173613 173712: gap of 100 bp  
173713 187445: contig of 13733 bp in length  
187446 187545: gap of 100 bp  
187546 191084: contig of 3539 bp in length  
191085 191184: gap of 100 bp  
191185 194709: contig of 3525 bp in length  
194710 194809: gap of 100 bp  
194810 197824: contig of 3015 bp in length  
197825 197924: gap of 100 bp  
197925 200623: contig of 2699 bp in length  
200624 200723: gap of 100 bp  
200724 203433: contig of 2710 bp in length  
203434 203533: gap of 100 bp  
203534 206523: contig of 2990 bp in length  
206524 206623: gap of 100 bp  
206624 208875: contig of 2252 bp in length  
208876 212659: gap of 100 bp  
212659 212659: contig of 3684 bp in length.  
Location/Qualifiers  
1. 212659  
/organism="Homo sapiens"

/db xref="taxon:9606"  
/chromosome="13"  
/clone="RP11-190P17"  
/clone\_1b="RP11-11.1"  
1. 48415  
/note="assembly fragment:03097  
fragment chain:1"  
misc\_feature  
48516..82643  
/note="assembly fragment:01394  
fragment chain:1"  
misc\_feature  
82744..89101  
/note="assembly fragment:01816  
fragment chain:1"  
misc\_feature  
89202..127367  
/note="assembly fragment:03032  
fragment chain:2"  
misc\_feature  
127468..169067  
/note="assembly fragment:00218  
fragment chain:2"  
misc\_feature  
169168..171259  
/note="assembly fragment:00015"  
misc\_feature  
171360..173612  
/note="assembly fragment:00477"  
misc\_feature  
173713..187445  
/note="assembly fragment:00780"  
misc\_feature  
187546..191084  
/note="assembly fragment:01634"  
misc\_feature  
191185..194709  
/note="assembly fragment:01673"  
misc\_feature  
194810..197824  
/note="assembly fragment:02018"  
misc\_feature  
197925..200623  
/note="assembly fragment:02519"  
misc\_feature  
200724..203433  
/note="assembly fragment:02839"  
misc\_feature  
203534..206523  
/note="assembly fragment:02989"  
misc\_feature  
206624..208875  
/note="assembly fragment:03934"  
misc\_feature  
208976..212659  
/note="assembly fragment:04063"  
BASE COUNT 61246 a 41295 c 42666 g 65947 t 1505 others  
ORIGIN  
Query Match 85.7%; Score 1442.8; DB 2; Length 212659;  
Best Local Similarity 94.1%; Pred. No. 0;  
Matches 1588; Conservative 87; Indels 13; Gaps 8;  
QY 1 ATGGCTACGAGAAATTTGACATGCTGATGCTGTTTACACAGAAAGTTC 120  
DB 83107 ATGGCTACGAGAAATTTGACATGCTGATGCTGTTTACACAGAAAGTTC 83166  
QY 61 GTATTCATTCAGAAATTTGACATGCTGATGCTGTTTACACAGAAAGTTC 120  
DB 83167 GTATTCATTCAGAAATTTGACATGCTGATGCTGTTTACACAGAAAGTTC 83226  
QY 121 GAAAACTAGATGAATTTAGCTTGACAGGCTAGTGCACATAGTGAATTAAGTGAAGA 180  
DB 83227 GAAAACTAGATGAATTTAGCTTGACAGGCTAGTGCACATAGTGAATTAAGTGAAGA 83286  
QY 181 GCATGGAAGCTTTAAAGAGTTCATGAAGAGCGCATGCGAGTCTTCAACAGTTT 240  
DB 83287 GCATGGAAGCTTTAAAGAGTTCATGAAGAGCGCATGCGAGTCTTCAACAG-TT 83345  
QY 241 AAAGCAGTATCTCTCATGTTGAGAACAAAGGCTTTTATGAGAGCATGAAG 300  
DB 83346 AAAGCAGTATCTCTCATGTTGAGAACAAAGGCTTTTATGAGAGCATGAAG 83405  
QY 301 ACTTACAGCAGAGAAAAACAGGGACCAAGTAGCAGACTTGAATTAAGACCAAGT 360  
DB 83406 CTTTATAGCAGAGAAAAACAGGGACCAAGTAGCAGACTTGAATTAAGACCAAGT 83465  
QY 361 GAGGCAAGATTAAAGCAGCTTTGAAAGAACAGGCTACACACTTGATGTGACTACAGGT 420



Db	83466	GAGGCAAAATTTAAGCACTCTTGAAAAGACAGCTTACACCTTGATGTGAOACTGGA	83525
Oy	421	CAGAGAACTATGAGAGACCACTTCAGATTCGGTTATTCAGCTCAGACGCTTCTGTT	480
Db	83526	CAGAGAAATATGAGAGACCACTTCAGATTCGGTTATTCAGCTCAGACGCTTCTGTT	83585
Oy	481	GGCACGTGATATTTGTGTGGGAAAGATCCCAAGATCTTTTGAGAGATGAGCTGTTC	540
Db	83586	GGCACGTGATATTTGTGTGGGAAAGATCCCAAGATCTTTTGAGAGATGAGCTGTTC	83645
Oy	541	TTATTTGAGAAAGCTGGAACCTATATGGATCTTCGTTATGATGATCCGCTCAGTGT	600
Db	83646	TTATTTGAGAAAGCTGCTGTCTTATATGGATCTTCGCTATATGATGATCCTCAGTGT	83705
Oy	601	CTCAA CAGAGTTATGCTGTTTGTCACTTTTGTGCAAAAGAACGACACAAAGGCTGTT	660
Db	83706	CTCAATATGAGTTATGCTGTTTGTGCACTTTTGTGCAAAAGACAGCTGGGAGGCTG--TTA	83766
Oy	661	AAACGTATATATATCATCATGAAATTCGTTCCGGGAACACATTTGTCTGTGATCTCAATT	720
Db	83764	AAACGTATATATATCATCATGAAATTCGTTCCGGGAACACATTTGTCTGTGATCTCAATT	83822
Oy	721	GCCAA CAAATAGGCTTTTGTGGGCTGATTTCCTAAG--TAAACCAAGAGCAGATTC	779
Db	83824	GCCAA CAAATAGGCTTTTGTGGGCTGATTTCCTAAGATTTAAACCAAGAGCAGATTC	83888
Oy	780	TGAGAAATTTAGC--AAAGTGA CAGAGGTTCTCAGATGTCAATTTTATACCAACCTG	838
Db	83884	TGAAGAAATTTTACAAAGATTA CAGAGGTTCTTACAGACGTCAATTTTATACCAACCG	83943
Oy	839	ATGACAAAGAAAAAAACAGAGCTTTTGTCTTGTAAATATGAAATTCACAAACGCTG	898
Db	83944	ATGACAAAGAAAAAAAC--AGGCTTTTGTCTTGTAAATATGAAATTCACAAACGCTG	84001
Oy	899	CCCAGGCAAGACGTAGGCTTAATGAGTGTAAAGTCAAAGCTGGGGAAATGTTGAACTG	958
Db	84002	CCCAGGCAAGGCGTATAGTTATGAGTGTAAAGTCAAAGCTGGGGAAATGTTGAACTG	84064
Oy	959	TTGAGTGGGCTGATCTCTATTTGAAGATCTGATCTTGAAGTTATGGCAAGTAAAGTGC	1018
Db	84062	TTGAAATGGGCTGATCTCTATTTGAAGATCTGATCTTGAAGTTATGGCAAGTAAAGTGC	84122
Oy	1019	TGTTGTATACGCACTTGTGCCACACGCTTA CAGAGAAATTTTGAAGAGCTCATTTAGTC	1078
Db	84122	TGTTGTATACGCACTTGTGCCACATCTGTAA CCGAAGAGATTTTGAAGAGCTCATTTAGTC	84181
Oy	1079	AGTTTGGGAAACTGGAACAGATTAAGACCTAAAGATTAATATGCTTCATTCATTTGATG	1138
Db	84182	AGTTTGGGAAACTGGAATAGATGAGAAAGTTAAAGATTAATCATTCATTCATTTGATG	84241
Oy	1139	AGAGATGTGCTGTCAAGGCTATGGAAGAAATGATGTTAAAGCTTGAGGGAGAA	1198
Db	84242	AGTGAATGTGCTGTCAAGGCTATGGAAGAAATGATGTTAAAGCTTGAGGGAGAA	84301
Oy	1199	ATATTGAATTTGTTTGTCTTAAGCCACAGATTA GAAAGAGAAAGAAAGAGCTCAGA	1258
Db	84302	ATATTGAATTTGTTTGTCCAAGCCCAAACTAGAAAGAGAAAGAA--AAAGGCTCAGA	84353
Oy	1259	GCCAGACACAAAGATCAAATGTATAGATGATTA CTAATTAATGTGCCACCTCATATGC	1318
Db	84360	GCCAGACACAAAGATCAAATGTATAGATGATTA CTAATTAATGTGCCACCTCATATGC	84419
Oy	1319	CTCCCCCAACAAGAGTCAAGGGCTGAGAGTGAAGTGGGCTATGATATCCCTCAGATT	1378
Db	84420	CCCCCTCCACACAGAGATGAGGGCCGAGAGGTAGAGTGTATATCTCTCCAGATT	84477
Oy	1379	ATTATGATACGAAGATTTATATGATTAATATGATTA TGAATTA CTAATCTATCTGTGT	1438
Db	84480	ATTATGATATGAAGATTTATATGATTAATATGATTA TGAATTA CTAATCTATCTGTGT	84533
Oy	1439	GATATGAATCCATATCTATGTGTTATGAAGATTTT CAATTTGAGCTTAAGAGAGGGGTG	1498

```

84540 GATATGAAGATCCACTATGATGTTATGAAGATTTTCAAGTTGAGGCTAGA-GAAGGGGTG 84598
QY 1499 GTAAGAGGAGCAAGGGGTGCTCTCTCATTCACAAGGTCGTGGGGCTGCTCTCCCGGTGTA 1558
Db 84599 GTAAGAGGAGCAAGGGGTGCTCTCTCATTCACAAGGTCGTGGGGCTGCTCTCCCGGTGTA 84658
QY 1559 GAGCGCGTTATTCACAGAGAGAGAGCCCTGATATAGCAAGAGGCGTTGCGCGGTGCGAGAG 1618
Db 84659 GAGCGCGTTATTCACAGAGAGAGAGCCTGATATAGCAAGAGGCGTTGCGAGGTGCGAGAG 84718
QY 1619 GAGGTGCCCAACAACAAGAGGCGCGCG--GGAAGAAAGGGGTGAGAGCGCGTCTGAC 1675
Db 84719 GAGGTGCCCAACAACAAGAGGCGCGCGGAGGAGAAAGGGGTGAGAGCGTCTGAC 84778
QY 1676 TGTTCACA 1683
Db 84779 TGTTCACA 84786
|||||
|||||

RESULT 11
BC032643
LOCUS
DEFINITION Homo sapiens, similar to NS1-associated protein 1, clone MGC:45213
IMAGE:5495201, mRNA, complete cds.
BC032643
ACCESSION BC032643
VERSION BC032643.1 GI:21619167
KEYWORDS MGC.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Strausberg,R.
TITLE Direct Submision
JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbe-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@hgrl.nih.gov
Aklter,N., Aylee,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakeslee,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R.,
Maduro,Q.L., Mastello,C., Maekel,B., Mastrian,S.D., McCloskey,J.C.,
McGowan,J., Pearson,R., Startitop,S., Thomas,P.J., Touchman,J.W.,
Tsurugicon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAX Plate: 69 Row: F Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5453805.
FEATURES
location/Qualifiers
1..3319
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:45213 IMAGE:5495201"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH MGC_67"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
98..1681

```

```
/codon_start=1
/product="Similar to NS1-associated protein 1"
/protein_id="AA32643.1"
/db_xref="GI:21619168"
/translation="MATEHYNKGTEPMDTTSVAVHSENFQTLIDAGLPQKAEKLD
ETVAELVAHSDLERALKEFNEDGALAVIQPKSDLSHVQKSAFLCGVMKTY
ROREKGTIVADSSSKGDEDAKALERGTLDVTGKRGPPDVSQCPVS
GTEIFVGIKPRDLFEDELVLPEKAGPIQDLRLMDPLGLNKGVAFTVCTEAAOE
AVELVNNHIEIRSGKHIGVCSIVANNELFVGSIPKSTKXOILBESPKTEGLTDVLY
HOPDDKKRGRGCELEVEDHKTAQYKVLFEVRLATVTEILERKPSFGKLERVK
LNDYAFIHDERDGAIVKAMEBNKQLEGENIEIVAKPDDKRRKKQKQAKNOM
YDDYVYGRPHMPPTRGKRGGRGYGPPDYEDYDYGVDYHNRGGEDEPY
YGRDFOVARGRGARGAAPSRRGAAPRGAGYSORGGPSARVGRARGAAG
OORGGOGGVEGPDILQ"
BASE COUNT      1055 a      569 c      712 g      983 t
ORIGIN
Query Match      81.2%; Score 1366.8; DB 9; Length 3319;
Best Local Similarity 90.5%; Pred. No. 5.5e-301;
Matches 1526; Conservative 0; Mismatches 52; Indels 108; Gaps 2;
QY 1 ATGGCTACAGAACTGTATGAAATGTACTGAAAGCCCATGATCTACTTCACGA 60
DB 98 ATGGCTACAGAACTGTATGAAATGTACTGAAAGCCCATGATCTACTTCACGA 157
QY 61 GTATTCATTCAGAAAATTTTCAGACATGCTGATGCTGTTTACAGAAAGTTGCT 120
DB 158 GTATTCATTCAGAAAATTTTCAGACATGCTGATGCTGTTTACAGAAAGTTGCT 217
QY 121 GAAAACTAGATGAAATTTTACGTTGACGAGCTAGTTGACATAGTGAATTAAGA 180
DB 218 GAAAACTAGATGAAATTTTACGTTGACGAGCTAGTTGACATAGTGAATTAAGA 277
QY 181 GCTATGGAAGCTTTTAAAGAGTTCAATGAAGACGCGCATTTGGCAGTTCACAGTTT 240
DB 278 GCTATGGAAGCTTTTAAAGAGTTCAATGAAGACGCGCATTTGGCAGTTCACAGTTT 337
QY 241 AAGACAGTATCTCTCTCATGTTTACAGAACAAAGTSCCTTTTATGAGAGTCATGAG 300
DB 338 AAGACAGTATCTCTCTCATGTTTACAGAACAAAGTSCCTTTTATGAGAGTCATGAG 397
QY 301 ACTTACAGGACAGAGAAAAACAGGGGACCAAGTAGCAGACTCTAGTAAAGACAGAT 360
DB 398 ACTTACAGGACAGAGAAAAACAGGGGACCAAGTAGCAGACTCTAGTAAAGACAGAT 457
QY 361 GAGGCAAAATTTAAGGACCTTTGGAAAGAACAGGCTACACACTTGATGACTACAGT 420
DB 458 GAGGCAAAATTTAAGGACCTTTGGAAAGAACAGGCTACACACTTGATGACTACAGT 517
QY 421 CAGAGAGATGAGAGGACCACTCCAGATTCGTTTATTCAGGTCAGAGCCCTTCGTT 480
DB 518 CAGAGAGATGAGAGGACCACTCCAGATTCGTTTATTCAGGTCAGAGCCCTTCGTT 577
QY 481 GGCAGTGAATTTTGTGGGAGAGATCCCAAGATCTGTTTGAAGTAGAGTTTCCA 540
DB 578 GGCAGTGAATTTTGTGGGAGAGATCCCAAGATCTGTTTGAAGTAGAGTTTCCA 637
QY 541 TTTATTTGAAGAGCTGACCTATATGAGGATCTTCTTTAATGATGATCCGCTCACTG 600
DB 638 TTTATTTGAAGAGCTGACCTATATGAGGATCTTCTTTAATGATGATCCGCTCACTG 697
QY 601 CTGAACAGAGTTATGCTGCTTCACCTTTTGAACAAAAGACAGACAGAGGCTGTT 660
DB 698 CTGAACAGAGTTATGCTGCTTCACCTTTTGAACAAAAGACAGAGGCTGTT 757
QY 661 AAATCTATATATCATGAAATTCGTTCCGGAAGACAGATTGCTGATCTCAGTT 720
DB 758 AAATCTATATATCATGAAATTCGTTCCGGAAGACAGATTGCTGATCTCAGTT 817
QY 721 GCCAACAATAGGCTTTTGTGGGCTGATCTCTAAGAGTAAACCAAGAGCAGATTTCT 780
DB 818 GCCAACAATAGGCTTTTGTGGGCTGATCTCTAAGAGTAAACCAAGAGCAGATTTCT 877
```

```
QY 781 GAGAAATTTGACAAAGTACAGAGGGCTCAGATGTCAATTTTATACCAACCACTGAT 840
DB 878 GAGAAATTTGACAAAGTACAGAGGGCTCAGATGTCAATTTTATACCAACCACTGAT 937
QY 841 GACAAAGAAAAAACAAGAGCTTTTCTTTCTTGAATATGAAATCAACAAACGCTGCC 900
DB 938 GACAAAGAAAAAACAAGAGCTTTTCTTTCTTGAATATGAAATCAACAAACGCTGCC 997
QY 901 CAGGCAAGAGTATGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 960
DB 998 C----- 998
QY 961 GAGTGGGCTGATCTTATGAAATCTGATCTGAAATATGCAAGGTAAGTAAAGTCTG 1020
DB 999 -----AGGTAAGTCTG 1012
QY 1021 TTTGTACGAACCTTGCCCAACAGGTAAACAGAAATTTTAAAGAAAGTAAAGTAAAG 1080
DB 1013 TTTGTACGAACCTTGCCCAACAGGTAAACAGAAATTTTAAAGAAAGTAAAGTAAAG 1072
QY 1081 TTTGGGAACTGGAACGAGTAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1140
DB 1073 TTTGGGAACTGGAACGAGTAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1132
QY 1141 AGAGATGGTCTGTCAAGGCTATGAAAGAAATGAATGTAAGTAAAGTAAAGTAAAGTAAAG 1200
DB 1133 CGAGATGGTCTGTCAAGGCTATGAAAGAAATGAATGTAAGTAAAGTAAAGTAAAGTAAAG 1192
QY 1201 ATGAAATTTGTTTCTTAAAGCAGCAGATCAGAAAGAGAAAGAAAGAAAGTCAAGAG 1260
DB 1193 ATGAAATTTGTTTCTTAAAGCAGCAGATCAGAAAGAGAAAGAAAGAAAGTCAAGAG 1252
QY 1261 CAAGCAGCAAGAAATCAATATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
DB 1253 CAAGCAGCAAGAAATCAATATGATGATGATGATGATGATGATGATGATGATGATGATG 1312
QY 1321 CCCCACAAAGAGGTGAGAGGCGTGAAGTAAAGGCTATGAGATATCTCCAGATTAAT 1380
DB 1313 CCCCACAAAGAGGTGAGAGGCGTGAAGTAAAGGCTATGAGATATCTCCAGATTAAT 1372
QY 1381 TATGATTAAGAAATTTATGATTTATGATTTATGATTTATGATTTATGATTTATGATTTAT 1440
DB 1373 TATGATTAAGAAATTTATGATTTATGATTTATGATTTATGATTTATGATTTATGATTTAT 1432
QY 1441 TATGAAGATCCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
DB 1433 TATGAAGATCCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1492
QY 1501 AAGAGAGCAAGGGTCTGCTCATCAGAGGTCGTGGGCTGCTCTCCCGGTGATA 1560
DB 1493 AAGAGAGCAAGGGTCTGCTCATCAGAGGTCGTGGGCTGCTCTCCCGGTGATA 1552
QY 1561 GCCGTTATTTCAAGAGAGAGGCTCTGATCAGAAAGGCGTTGCGGCTGAGAGGA 1620
DB 1553 GCCGTTATTTCAAGAGAGAGGCTCTGATCAGAAAGGCGTTGAGGTGAGAGGA 1612
QY 1621 GGTGCCCAACAAAGAGGCGCGG---GGAAAAAGGGGTGAGAGGCGGCTCTGACCTG 1677
DB 1613 GGTGCCCAACAAAGAGGCGCGGCGGAGGAAAAAGGGGTGAGAGGCGGCTCTGACCTG 1672
QY 1678 TTAACA 1683
DB 1673 TTAACA 1678
RESULT 12
LOCUS AY034482 1767 bp mRNA linear PRI 03-OCT-2001
DEFINITION Homo sapiens hnrnp Q2 mRNA, complete cds.
ACCESSION AY034482
VERSION AY034482.1 GI:15809587
KEYWORDS SOURCE Homo sapiens.
```

## ORGANISM

Homo sapiens

## REFERENCE

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 1767)  
Mouratiolos, Z., Abel, L., Yong, J., Kataoka, N. and Dreyfuss, G.  
SNR interacts with a novel family of hnRNP and spliceosomal  
proteins

## JOURNAL

EMBO J. 20 (19), 5443-5452 (2001)

## MEDLINE

21458434

## PUBMED

11574476

## REFERENCE

2 (bases 1 to 1767)  
Mouratiolos, Z.  
Direct Submission

## JOURNAL

Submitted (11-MAY-2001) HMI/ Biochemistry, University of  
Pennsylvania, 330 CRB, 415 Curie Blvd., Philadelphia, PA 19104, USA

## FEATURES

Location/Qualifiers

## CDS

1..1767  
/db\_xref="taxon:9606"  
/codon\_start=1  
/product="hnRNP Q2"  
/protein\_id="AAK59704.1"  
/db\_xref="GI:15809588"  
/translation="MATEHYNGTEEPMDTSVISHENFOTLLDAGLPQVAKLKD  
EIVVAGLVASHSDDEPAIEMKEPNEGALAVIQPDPDLSHONKSAFLGCMKTD  
RQREKQSTKVDSDSKDEPKIKALLERTYTDVTTGQRKGPPDSVYSGQSSV  
GTEIPVKGIPRDLPEDELVPLFERAGIPWLRIMDPLGLANGYAVVCTEAAOE  
AVKLNNHNRISGKHIGVCI SVANNRLFVSIIPSKKEIILEFSAVTEGLDVLILY  
HOPDDKKKNSPCFLEYEDHKTAQVAVLVFNINAKVTEILEKFAVQGLERVYK  
LKDYAFIHPEBDGA VKAMEENMGKDLLEGNIIEVFAKPDOKREKKAQRAKQNM  
YDYVYVGPHPMPPTGRGRGGGAGVDPDYGVYDYDYGVYDHYNRGGYEDPY  
KYVEDPVGARGRGARGARGAABSRGGAPPRGRAGYSGRGSGSRGTRARGAQ  
QQRGVRGARGRGARGRGARGRGADGNQDPDSRRQTNQNMWSQPLAQPLQSGDHS  
GNVGYSENDEFQDTFGQWK"BASE COUNT 561 a 310 c 453 g 443 t  
ORIGINQuery Match 79.8%; Score 1343.4; DB 9; Length 1767;  
Best Local Similarity 89.8%; Pred. No. 1.1e-295;  
Matches 1499; Conservative 0; Mismatches 66; Indels 105; Gaps 1;

QY 1 ATGGCTACAGAAACATGTTAATGAAATGTTACTGAAGGCCCATGATCTACTTACAGA 60  
Db 1 ATGGCTACAGAAACATGTTAATGAAATGTTACTGAAGGCCCATGATCTACTTACAGA 60

QY 61 GTTATCCATTGAGAAATTTTCAAGCATTTGCTGTGATTGCTTACACAGAAAGTTGCT 120  
Db 61 GTTATCCATTGAGAAATTTTCAAGCATTTGCTGTGATTGCTTACACAGAAAGTTGCT 120

QY 121 GAAAACTAGATGAATTTAGCTTGCAGAGGCTAGTTGCACATAGATTGATGAGAAAGA 180  
Db 121 GAAAACTAGATGAATTTAGCTTGCAGAGGCTAGTTGCACATAGATTGATGAGAAAGA 180

QY 181 GCTATCGAAGCTTTAAAGAGTTCAATGAGACGGCGATTTGGCAGTCTTCAACAGTTT 240  
Db 181 GCTATCGAAGCTTTAAAGAGTTCAATGAGACGGCGATTTGGCAGTCTTCAACAGTTT 240

QY 241 AAGACAGTGTCTCTCATGTTTCAAGCAAAAGTGCCTTTTATGTGAGTCAATGAG 300  
Db 241 AAGACAGTGTCTCTCATGTTTCAAGCAAAAGTGCCTTTTATGTGAGTCAATGAG 300

QY 301 ACTTACAGGAGAGAGAAAAAAGAGGACCAAGTACAGACTCTAGTAAAGACAGAT 360  
Db 301 ACTTACAGGAGAGAGAAAAAAGAGGACCAAGTACAGACTCTAGTAAAGACAGAT 360

QY 361 GAGGCAAAAGATTTAAGGCACTTTTGAAGAAACAGGCTACACACTTGATGTGACTACAG 420  
Db 361 GAGGCAAAAGATTTAAGGCACTTTTGAAGAAACAGGCTACACACTTGATGTGACTACAG 420

QY 421 GAGGAGAGATGAGAGCACTTCACAGATTCGTTTATTCAGGTCAAGCAGCTTCTGT 480  
Db 421 GAGGAGAGATGAGAGCACTTCACAGATTCGTTTATTCAGGTCAAGCAGCTTCTGT 480

QY 481 GGCACTGAGATATTTGNGGGGAAGATCCCAAGATCTGTTGAGATGAGCTTTTCCA 540  
Db 481 GGCACTGAGATATTTGNGGGGAAGATCCCAAGATCTGTTGAGATGAGCTTTTCCA 540

QY 541 TTATTTGAGAAAGCTGAGACCTATATGAGATCTTCTGTTTATGATGATCCGCTACATG 600  
Db 541 TTATTTGAGAAAGCTGAGACCTATATGAGATCTTCTGTTTATGATGATCCGCTACATG 600

QY 601 CTCAACAGAGATTATGCTTTTGTCACTTTTGTACAAAGAGACAGACAGAGCTGTT 660  
Db 601 CTCAACAGAGATTATGCTTTTGTCACTTTTGTACAAAGAGAGCTGAGAGCTGTT 660

QY 661 AAACGTATATATATATGATGAAATTCGTCGGGAGAGACATTTGTCGTCATCTCA 720  
Db 661 AAACGTATATATATATGATGAAATTCGTCGGGAGAGACATTTGTCGTCATCTCA 720

QY 721 GCCAACAATAGGCTTTTGTGGGCTCGATTCCTAAGATGAAACCAAGAGACAGATCTT 780  
Db 721 GCCAACAATAGGCTTTTGTGGGCTCGATTCCTAAGATGAAACCAAGAGACAGATCTT 780

QY 781 GAGGAATTTAGCAAGAGTACAGAGGCTCTCAAGATGTCAATTTATACCAACACTGAT 840  
Db 781 GAGGAATTTAGCAAGAGTACAGAGGCTCTCAAGATGTCAATTTATACCAACACTGAT 840

QY 841 GACAAAGAAAAAACAAGAGCTTTTGTCTTCTGATATGAAATGACAAACAGCTGCC 900  
Db 841 GACAAAGAAAAAACAAGAGCTTTTGTCTTCTGATATGAAATGACAAACAGCTGCC 900

QY 901 CAGGCAAGAGTATAGGCTAATAGTGTGAAAGTCAAGTCTGGGAAATGTGAACTGTT 960  
Db 901 C----- 901

QY 961 GAGTGGCTGATCTTATGAAATCTGATCTGAAAGTATGCAAGTAAAGTCTG 1020  
Db 961 -----AGTAAAGTCTG 915

QY 1021 TTGTGACGCACTTGTGCAACAGGTAACAGAAATTTAGAAAGTATTAATGACAG 1080  
Db 916 TTGTGACGCACTTGTGCAACAGGTAACAGAAATTTAGAAAGTATTAATGACAG 975

QY 1081 TTGGGAACTGAAACGAGTAAGAAAGCTAAAGATTAATGCTTTCATCTTTTATGAG 1140  
Db 976 TTGGGAACTGAAACGAGTAAGAAAGCTAAAGATTAATGCTTTCATCTTTTATGAG 1035

QY 1141 AGAGATGCTGCTGCAAGGCTATGAAAGAAATGAAATGTAAGACTTGAAGGAGAAAT 1200  
Db 1036 CGAGATGCTGCTGCAAGGCTATGAAAGAAATGAAATGCAAGACTTGAAGGAGAAAT 1095

QY 1201 ATTGAATTTGTTTGTATAGCCACCAAGATCAGAAAGGAAAGAAAGCTCAGAGG 1260  
Db 1096 ATTGAATTTGTTTGTATAGCCACCAAGATCAGAAAGGAAAGAAAGCTCAGAGG 1155

QY 1261 CAAGCAGCAAGATCAATATGATGATGATTAATTAATGATGATGATGATGATGATGAT 1320  
Db 1156 CAAGCAGCAAGATCAATATGATGATGATTAATTAATGATGATGATGATGATGATGAT 1215

QY 1321 CCCCACCAAGAGTGCAGGCGGTGAGAGTGAAGTGGCTATGATATCTCCAGATTAAT 1380  
Db 1216 CCCCACCAAGAGTGCAGGCGGTGAGAGTGAAGTGGCTATGATATCTCCAGATTAAT 1275

QY 1381 TATGATATGAGAAATTTATGATATGATATGATATGATATGATATGATATGATATG 1440  
Db 1276 TATGATATGAGAAATTTATGATATGATATGATATGATATGATATGATATGATATG 1335

QY 1441 TATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 1500  
Db 1336 TATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 1395

QY 1501 AGAGAGCAAGGAGTGTGCTCAGATCCAGAGTGTGAGGCTGCTCCCGTGTGAG 1560  
Db 1396 AGAGAGCAAGGAGTGTGCTCAGATCCAGAGTGTGAGGCTGCTCCCGGAGTGA 1455

QY	1561	GC	CG	GT	AT	TC	CA	GA	GA	GA	GC	CC	TC	GA	TG	AT	CA	GA	GA	GG	GT	TC	CG	CG	TC	CA	GA	GA	1620
DB	1456	GC <td>CG <td>GT <td>AT <td>TC <td>CA <td>GA <td>GA <td>GA <td>GC <td>CC <td>TC <td>GA <td>TG <td>AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CG <td>GT <td>AT <td>TC <td>CA <td>GA <td>GA <td>GA <td>GC <td>CC <td>TC <td>GA <td>TG <td>AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GT <td>AT <td>TC <td>CA <td>GA <td>GA <td>GA <td>GC <td>CC <td>TC <td>GA <td>TG <td>AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	AT <td>TC <td>CA <td>GA <td>GA <td>GA <td>GC <td>CC <td>TC <td>GA <td>TG <td>AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	TC <td>CA <td>GA <td>GA <td>GA <td>GC <td>CC <td>TC <td>GA <td>TG <td>AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CA <td>GA <td>GA <td>GA <td>GC <td>CC <td>TC <td>GA <td>TG <td>AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GA <td>GA <td>GA <td>GC <td>CC <td>TC <td>GA <td>TG <td>AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GA <td>GA <td>GC <td>CC <td>TC <td>GA <td>TG <td>AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GA <td>GC <td>CC <td>TC <td>GA <td>TG <td>AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GC <td>CC <td>TC <td>GA <td>TG <td>AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CC <td>TC <td>GA <td>TG <td>AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	TC <td>GA <td>TG <td>AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GA <td>TG <td>AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td>	TG <td>AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td></td>	AT <td>CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td></td>	CA <td>GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td></td>	GA <td>GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td></td>	GA <td>GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td></td>	GG <td>GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td></td>	GT <td>TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td></td>	TC <td>CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td></td>	CG <td>CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td></td>	CG <td>TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td></td>	TC <td>CA <td>GA <td>GA <td>1515</td> </td></td></td>	CA <td>GA <td>GA <td>1515</td> </td></td>	GA <td>GA <td>1515</td> </td>	GA <td>1515</td>	1515
QY	1621	GG	TG	CC	CA	CA	CA	AA	GA	GA	GC <td>CG <td>CG <td>GG <td>GA</td> <td>AA</td> <td>GA</td> <td>GG</td> <td>GC</td> <td>TG <td>CA</td> <td>GC <td>CG <td>GC <td>GC <td>TC <td>1669</td> <td></td> <td></td> </td></td></td></td></td></td></td></td></td>	CG <td>CG <td>GG <td>GA</td> <td>AA</td> <td>GA</td> <td>GG</td> <td>GC</td> <td>TG <td>CA</td> <td>GC <td>CG <td>GC <td>GC <td>TC <td>1669</td> <td></td> <td></td> </td></td></td></td></td></td></td></td>	CG <td>GG <td>GA</td> <td>AA</td> <td>GA</td> <td>GG</td> <td>GC</td> <td>TG <td>CA</td> <td>GC <td>CG <td>GC <td>GC <td>TC <td>1669</td> <td></td> <td></td> </td></td></td></td></td></td></td>	GG <td>GA</td> <td>AA</td> <td>GA</td> <td>GG</td> <td>GC</td> <td>TG <td>CA</td> <td>GC <td>CG <td>GC <td>GC <td>TC <td>1669</td> <td></td> <td></td> </td></td></td></td></td></td>	GA	AA	GA	GG	GC	TG <td>CA</td> <td>GC <td>CG <td>GC <td>GC <td>TC <td>1669</td> <td></td> <td></td> </td></td></td></td></td>	CA	GC <td>CG <td>GC <td>GC <td>TC <td>1669</td> <td></td> <td></td> </td></td></td></td>	CG <td>GC <td>GC <td>TC <td>1669</td> <td></td> <td></td> </td></td></td>	GC <td>GC <td>TC <td>1669</td> <td></td> <td></td> </td></td>	GC <td>TC <td>1669</td> <td></td> <td></td> </td>	TC <td>1669</td> <td></td> <td></td>	1669		
DB	1516	GG	TG	CC	CA	CA	CA	AA	GA	GA	GC <td>CG <td>CG <td>GG</td> <td>GA</td> <td>AA</td> <td>GA</td> <td>GG</td> <td>GC <td>TG <td>CA</td> <td>GC <td>CG <td>GC <td>GC <td>TC <td>1564</td> <td></td> <td></td> </td></td></td></td></td></td></td></td></td>	CG <td>CG <td>GG</td> <td>GA</td> <td>AA</td> <td>GA</td> <td>GG</td> <td>GC <td>TG <td>CA</td> <td>GC <td>CG <td>GC <td>GC <td>TC <td>1564</td> <td></td> <td></td> </td></td></td></td></td></td></td></td>	CG <td>GG</td> <td>GA</td> <td>AA</td> <td>GA</td> <td>GG</td> <td>GC <td>TG <td>CA</td> <td>GC <td>CG <td>GC <td>GC <td>TC <td>1564</td> <td></td> <td></td> </td></td></td></td></td></td></td>	GG	GA	AA	GA	GG	GC <td>TG <td>CA</td> <td>GC <td>CG <td>GC <td>GC <td>TC <td>1564</td> <td></td> <td></td> </td></td></td></td></td></td>	TG <td>CA</td> <td>GC <td>CG <td>GC <td>GC <td>TC <td>1564</td> <td></td> <td></td> </td></td></td></td></td>	CA	GC <td>CG <td>GC <td>GC <td>TC <td>1564</td> <td></td> <td></td> </td></td></td></td>	CG <td>GC <td>GC <td>TC <td>1564</td> <td></td> <td></td> </td></td></td>	GC <td>GC <td>TC <td>1564</td> <td></td> <td></td> </td></td>	GC <td>TC <td>1564</td> <td></td> <td></td> </td>	TC <td>1564</td> <td></td> <td></td>	1564		
<p>RESULT 13</p> <p>AX011753 3101 bp DNA linear PAT 06-SEP-2000</p> <p>AX011753</p> <p>DEFINITION Sequence 156 from Patent W0995858.</p> <p>LOCUS AX011753</p> <p>VERSION AX011753.1 GI:9998277</p> <p>KEYWORDS</p> <p>SOURCE human.</p> <p>ORGANISM Homo sapiens</p> <p>REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (baes 1 to 3101)</p> <p>AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and Pilarsky,C.</p> <p>TITLE Human nucleic acid sequences obtained from pancreas tumor tissue</p> <p>JOURNAL Patent: WO 995858-A 156 04-NOV-1999.</p> <p>SCHMITT ADMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN BERND (DE); ROSENTHAL ANDRE (DE); METZGEN GRS FUER GENOMFORSCHUN (DE); PILARSKY CHRISTIAN (DE)</p> <p>FEATURES</p> <p>source 1..3101</p> <p>location/Qualifiers</p> <p>1..3101 /organism="Homo sapiens"</p> <p>BASE COUNT 980 a 540 c 670 g 911 t</p> <p>ORIGIN</p>																													
<p>Query Match 72.7% Score 1222.8; DB 6; Length 3101;</p> <p>Best Local Similarity 89.6% Pred. No. 3.7e-268;</p> <p>Matches 1382; Conservative 0; Mismatches 52; Indels 108; Gaps 2;</p>																													
QY	145	GC	AG	GC	TAG	TGC	CA	TAG	TGA	TG	AT	TGA	TGA	AA	GA	GC	TAT	GA	AA	GA	GC	TAT	TA	AA	GA	GC	TTT	CA	204
DB	86	GC	AG	GC	TAG	TGC	CA	TAG	TGA	TG	AT	TGA	TGA	AA	GA	GC	TAT	GA	AA	GA	GC	TAT	TA	AA	GA	GC	TTT	CA	145
QY	205	AAT	GAA	AG	CG	CG	CAT	TGC	CA	GT	GCT	CTT	CA	CA	AGT	TAA	GA	CA	GT	AT	CT	CT	CA	TG	TT	264			
DB	146	AAT	GAA	AG	CG	CG	CAT	TGC	CA	GT	GCT	CTT	CA	CA	AGT	TAA	GA	CA	GT	AT	CT	CT	CA	TG	TT	205			
QY	265	CAG	A	CA	AA	AG	GC	CTT	TAT	TG	AG	TC	AT	GA	AA	GA	CA	CTT	TA	CA	GC	CA	GA	GA	AA	AA	CA	324	
DB	206	CAG	A	CA	AA	AG	GC	CTT	TAT	TG	AG	TC	AT	GA	AA	GA	CA	CTT	TA	CA	GC	CA	GA	GA	AA	AA	CA	265	
QY	325	GG	GA	CC	AA	AG	TAG	CAG	ACT	CT	AG	TAA	AG	CA	CA	GA	TAG	AG	CA	AA	GA	TT	TA	AG	CA	CTT	TTG	384	
DB	266	GG	GA	CC	AA	AG	TAG	CAG	ACT	CT	AG	TAA	AG	CA	CA	GA	TAG	AG	CA	AA	GA	TT	TA	AG	CA	CTT	TTG	325	
QY	385	GAA	A	GA	AA	CG	GCT	CA	CA	CTT	AT	G	TG	CA	CA	GT	CA	GA	AG	T	GA	GA	CA	CA	CA	CT	444		
DB	326	GAA	A	GA	AA	CG	GCT	CA	CA	CTT	AT	G	TG	CA	CA	GT	CA	GA	AG	T	GA	GA	CA	CA	CA	CT	385		
QY	445	CG	AG	AT	CC	GT	TAT	T	CA	AG	TG	CA	GA	GC	CC	TT	CG	TT	GA	CA	CT	GA	AT	TT	TG	GG	GA	504	
DB	386	CG	AG	AT	CC	GT	TAT	T	CA	AG	TG	CA	GA	GC	CC	TT	CG	TT	GA	CA	CT	GA	AT	TT	TG	GG	GA	445	
QY	505																												

D	566	ACITTTTGTCACAAAAGAGACGCTCAGAGAGCGCTGTTAACTGTATATATATATGAAATT	625
O	685	CGTTCGCGGAGACATTTGCTGTCTGCATCTCAGTTTGCCAAACAATAGCGTTTTGTGCGC	744
D	626	CGTTCGTGAAAACATATTTGCTGTCTGCATCTCAGTTTGCCAAACAATAGCGTTTTGTGCGC	685
O	745	TCGATTCCTTAAGGTAAACCAAGAGACGATTTCTTGAGAAATTTAGCAAACTGACAG	804
D	686	TCATTCCTTAAGGTAAACCAAGAGACGATTTCTTGAGAAATTTAGCAAACTGACAG	745
O	805	GGCTCACAGATGTCATTTATATACCAACCTGATGACAGAAAAAAAACAGAGCTTT	864
D	746	GGCTCTACAGACGTCATTTATATACCAACCGGATGACAGAAAAAAAACAGAGCTTT	805
O	865	TGCTTTCTTGAAATGGAAGTCACAAAACAGCTGCCAGAGCAAGAGTAGCTAATGAT	924
D	806	TGCTTTCTTGAAATGGAAGTCACAAAACAGCTGCC-----	842
O	925	GGTAAAGTCAAAAGCTGCGGGAAATGTTGGAAGCTGTTGAAGTGGGCTGATCTATGGAAT	984
D	843	-----	842
O	985	CCTGATCCTGAAGTTATGCGAAAAGTAAAGCTGTGTTGTAACGAACTTGCCCAACG	1044
D	843	-----AGTTAAAGTGTCTGTTGTAGCAACTTGCCAAATCT	880
O	1045	GTAACAGAGAAATTTTGAAGAAAGTCATTTAGTCAGTTTGAGAACTGGACGATGAAG	1104
D	881	GTAACAGAGAAATTTTGAAGAAAGTCATTTAGTCAGTTTGAGAACTGGAAAGATGAAG	940
O	1105	AAGCTAAAGATTTAGCTTTTCATTTGATGAGAGAGATGGTGTCTCAAGCTATG	1164
D	941	AAGTTAAAGATTTAGGTCATTTCAATTTGATGAGACGAGATGGTGTCTCAAGCTATG	1000
O	1165	GAGAAATGATGTGTAAAGCTTGAGGGGAGAAATATGAAATGTTTTGCTTAAGCA	1222
D	1001	GAGAAATGATGTGTAAAGCTTGAGGGGAGAAATATGAAATGTTTTGCTTAAGCA	1060
O	1225	CCAGATCAGAGAGAGAAAGAAAGAAAGCTCAGAGGCAAGCAGCAAGATCAATGTAT	1284
D	1061	CCAGATCAGAGAGAGAGAAAGAAAGAAAGCTCAGAGGCAAGCAGCAAGATCAATGTAT	1122
O	1285	GATGATTACTATATATGCTCACTCATATGCTCTCCCAACAAGAGTGAAGGCGT	1344
D	1121	GACGATTACTATATATGCTCACTCATATGCTCTCCCAACAAGAGTGAAGGCGT	1180
O	1345	GGAAGTAGAGGTGCTATGATATCCTCCAGATATATATGATAGGAATATATATAT	1400
D	1181	GGAAGTAGAGGTGCTATGATATCCTCCAGATATATATGATAGGAATATATATAT	1240
O	1405	TATATGTTATGATTAACAATACTATATCGGTGGATGATGAAATCCATATAGTTAT	1464
D	1241	TATATGTTATGATTAACAATACTATATCGGTGGATGATGAAATCCATATAGTTAT	1300
O	1465	GAAAGTTTTCAAGTTGAGGCTAAGAGAGAGAGGAGTGTAGAGAGCAAGGAGTCTGCTCA	1524
D	1301	GAAAGTTTTCAAGTTGAGGCTAAGAGAGAGAGGAGTGTAGAGAGCAAGGAGTCTGCTCA	1360
O	1525	TTCCAGAGGTGCTGGGCTGCTCTCCCGGTGTATGAGCCGGTATATTCACAGAGAGAGC	1588
D	1361	TTCCAGAGGTGCTGGGCTGCTCTCCCGGTGTATGAGCCGGTATATTCACAGAGAGAGT	1420
O	1585	CCTGATCAGCAAGAGCGCTTGCGCGGTGCGAGAGAGTGGCCCAACAACAAGAGAGCGC	1644
D	1421	CCTGATCAGCAAGAGCGCTTGCGCGGTGCGAGAGAGTGGCCCAACAACAAGAGAGCGC	1480
O	1645	GG---GGGAAAAGGGGTCGAGGCGCGGTCTGACCTGTTCAA	1683
D	1481	GGGCAAGGAAAAAGGGGTCGAGGCGCGGTCTGACCTGTTCAA	1522

LOCUS AF41128 1899 bp mRNA linear ROD 04-JAN-2002  
 DEFINITION Mus musculus CD-1 heterogeneous nuclear ribonucleoprotein R mRNA,  
 complete cds.  
 ACCESSION AF41128  
 VERSION AF41128.1 GI:17066598  
 KEYWORDS  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1899)  
 Roscoli, W., Kronting, A.K., Ohndorf, U.M., Steegborn, C., Jablonka, S.  
 and Sendtner, M.  
 TITLE Specific interaction of Smn, the spinal muscular atrophy  
 determining gene product, with hnRNP-R and gry-rbp/hnRNP-Q: a role  
 for Smn in RNA processing in motor axons?  
 Hum. Mol. Genet. 11 (1), 93-105 (2002)  
 JOURNAL Hum. Mol. Genet. 11 (1), 93-105 (2002)  
 MEDLINE 21634625  
 PUBMED 11773003  
 REFERENCE 2 (bases 1 to 1899)  
 Roscoli, W., Kronting, A.-K. and Sendtner, M.  
 TITLE Direct Submission  
 Submitted (30-OCT-2001) Inst. f. Clinical Neurobiology, Muerzburg  
 University, Josef-Schneider Str. 11, Muerzburg 97080, Germany  
 FEATURES  
 source  
 1. 1899  
 /organism="Mus musculus"  
 /strain="CD-1"  
 /db\_xref="taxon:10090"  
 1. 1899  
 /note="hnRNP-R"  
 /codon\_start=1  
 /product="heterogeneous nuclear ribonucleoprotein R"  
 /protein\_id="AA15332.1"  
 /db\_xref="GI:17066598"  
 /translation="MANQVNGNAVOLKEESEPMTSSVTHTEHYKTLIEAGLPQKVAE  
 RLDEIFQTVLAAYVDLDERAIDALRENEBGLSVLQPFESDLSHYONSAFLGVN  
 KTYRREKSGSKVQESTKPEDEAKIKALIRTYGLTVTGGORVGGPPDSVSGVO  
 PGCTEYFNKIPKIPRLYDEDELYPIPERAGPIWDLRLMDPLSGNRRYATITGGKKA  
 AOEAVKLCDSLEIRBGRALGVCSVANNRLVGSIPNKKENTLEFSSVYTBELVDV  
 ILVHPDPPDKRRNRCFLLEYEDHNSAQAARRRLMSGVKVMVNVYEMADPVEEDP  
 EVNAKVLFLVRNLTATVTEBILKESFSEKGLSKVDYAVFHEDEGAALKAMD  
 EMNGKEIEGEIEIYLAKPPDKKREKRAQASRSTAYEDYVHPPPRPPRGG  
 RGGKGGYGPDPYEDYDDYDYDYDHYDRGGEDPYGYDDYGVAVRRGGRRGR  
 GAPPREGARPPRGARAGYSORCAPIGPPRGSGGGGAPQQRGSGSGAGNRGSG  
 NVGKKRADGTNPDSKRRKOTNMQNMGSOP1AQDPLQGGSDYSNGYNNDDQEFYQ  
 DTYGQQWK"

BASE COUNT 581 a 362 c 548 g 408 t  
 ORIGIN  
 Query Match 53.2%; Score 896; DB 10; Length 1899;  
 Best Local Similarity 73.6%; Pred. No. 1,2e-193;  
 Matches 118; Conservative 0; Mismatches 415; Indels 9; Gaps 3;

QY 51 TACTTCAGCAGTTATCCATTGAGAAATTTTCAGACATTGCTTGATGCTTACCA 110  
 |||||  
 DB 60 TACTTCAGGTTATCTACACAGAACTACAGACACTGATAGGAGGCTCCCA 119  
 |||||  
 QY 111 GAAAGTTGCTGAAAAAATAGATGAATTTACGTTGACAGGCTAGTTGACAT 170  
 |||||  
 DB 120 GAAAGTTGCTGAAAAAATAGATGAATTTACAGAGGATGTTAGCTTATG 179  
 |||||  
 QY 171 AGATGAAGAGCTATCCGAACTTTTAAAGATTCAATGAAGCGCGCATTTGG 230  
 |||||  
 DB 180 TGAATGAAGAGCTATCCGAACTTTTAAAGATTCAATGAAGAGCTTGTCT 239  
 |||||  
 QY 231 TCAACAGTTTAAAGCAGTATCTCTCATGTTCAAGACAAAGTGCCTTTTATG 290  
 |||||  
 DB 240 GCAAGAGTTTAAAGAAATGACTCTCTCATGTTCAAGACAAAGTGCCTTT 299  
 |||||  
 QY 291 AGTCATGAAGCTTACAGGACAGAGAAAAACAGGGGACCAAGTAGACACTTA 350  
 |||||  
 DB 300 AGTTATGAAGACTTACAGGACAGAGAAAAACAGGACAGAGTAGTCAAGAGT 359  
 |||||

QY 351 AGAACCATGATGAGCAAAAGATTAAAGCACTTTGAGAAAAGACGCTACACTT 410  
 |||||  
 DB 360 AGGGCCCGATGAAAGCAAAATTCAGAGCTTGTGAGAGAGACTGTATATCT 419  
 |||||  
 QY 411 GACTACAGGTCAAGAGAAATGATGAGAGACCACTCCAGATTCCGTTATTCAG 470  
 |||||  
 DB 420 GACCAAGGACAGAGAAATGATGAGAGCTCCCAACAGAGTGTACTCGGATG 479  
 |||||  
 QY 471 GCCTTGTGTGGCACTGAGATATTGTGGGAAAGATCCCAAGATCTGTTGAG 530  
 |||||  
 DB 480 GCCTGAGATTGGACAGAGGCTCTTCGTAAGTAAATCCAAAGACTTGTATG 539  
 |||||  
 QY 531 GCTTGTCCATTTATTTGAGAAAGCTGACCTTATATGAGATCTTCCTTAT 590  
 |||||  
 DB 540 ATTATGATCCCTTTTGAAGAGGCTGTCTATCTGGGATCTACGGCTCATG 599  
 |||||  
 QY 591 GCTCACTGTCTCAACAGAGTTATGCGTTGTCACTTTTGTACAAAGAACAG 650  
 |||||  
 DB 600 ACTCTCTGTGACAGACAGAGGCTATGATTCATCACTTCTGTGAAAAGAG 659  
 |||||  
 QY 651 AGAGGCTGTAAACCTGATATATATCATGAAATTCGTCGGAGACATTTG 710  
 |||||  
 DB 660 GAAAGCAGTTAAGCTGTGTACAGACTATGAGATTGCGCTGTAAACACT 719  
 |||||  
 QY 711 CATCTCAGTTGCCAACAATAGGCTTTTGTGGGCTCGATTCCTTAAGATTA 770  
 |||||  
 DB 720 CATTTCTGTGGCAACAGACCTTTTGTGATTCATTCACAAAGATTAAGA 779  
 |||||  
 QY 771 GCAGATCTTGAGGAATTTGACAAAGTACAGAGGCTCTCAGATGTCAATTT 830  
 |||||  
 DB 780 GAACATTTCTGAAAGAGTTCACTAATGACAGAGGTTTGTGACCTTATCT 839  
 |||||  
 QY 831 CCAACTGATGACAGAAAAAACAAGGCTTTTGTCTTGAATATGAGATCA 890  
 |||||  
 DB 840 TCAACCCGATGACAAAAAAGAAATCGGGGCTTCTGCTTCTTGAATATG 899  
 |||||  
 QY 891 AAGCGTCCCAAGCAGACAGTACGCTTAATGATGTGTAAGTCAAACTCG 950  
 |||||  
 DB 900 GTCCAGCAGCAACAGCCAGCGCGCTGATGATGAAAGTCAAAATGTG 959  
 |||||  
 QY 951 TGCAACTGTGAGTGGGCTGATCTTATTAAGATCTCGATCTGAAGTTAG 1010  
 |||||  
 DB 960 GGTTAAGTAGAGTGGGCTGATCCCGTGGAGAGCCGATCCAGAACTCAT 1019  
 |||||  
 QY 1011 AAAAGTCTGTTGTACGCAACCTTGCACACGATACAGAAATTTTAAAG 1070  
 |||||  
 DB 1020 AAAAGTTTATTGTGAGAAACTTGGCACTACGATGACAGAAATTTTAA 1079  
 |||||  
 QY 1071 ATTATGACATTTGGGAAACTGACAGAGTAAAGACTTAAAGATTATGCT 1130  
 |||||  
 DB 1080 ATTATGACATTTGGGAAACTGACAGAGTAAAGACTTAAAGATTATGCT 1139  
 |||||  
 QY 1131 TTTTATGAGAGAGATGCTGCTCAAGCTATGAGAAATGATGTAAGCTTGA 1190  
 |||||  
 DB 1140 TTTTGAAGACAGGAGCTGCTGTTAAGCTATGAGAAATGATGAAAGAA 1199  
 |||||  
 QY 1191 GGGAGAAATATTGAAATTTGTTTGTAAACCAACGATCAGAAAGGAAAGAA 1250  
 |||||  
 DB 1200 AGGGAGAAATTTGAAATTTGTTTGTAAACCAACGATCAGAAAGGAG 1259  
 |||||  
 QY 1251 AGCTCAGAGCAGACAGCAAGAAATCAAAATGATATGATTTACTATATG 1310  
 |||||  
 DB 1260 AGCTCAGACAGGCTCTCAGAAAGCACTGGATGAAATTTATTTATCA 1319  
 |||||  
 QY 1311 TCATATGCTCTCCCAACAGAGGTCAGAGGCTGAGAGTATGAGTGTATG 1370  
 |||||  
 DB 1320 CCGAATGCACTTCATATGAGAGGCGAGGCTGTGAGGAGAGTGAATGCT 1379  
 |||||  
 QY 1371 TCCAGATTTATTTATGATACGAG---ATTATATGATTTATTTATGATTA 1427  
 |||||  
 DB 1380 TCCAGATTTATTTATGATACGACTATGATGATTTACTATGATTTATTA 1439  
 |||||

QY	1428	CHATGCTGGTGAATATGAAGATCATCATGATGTTATGAAGT---	TTTCAAGTTGGAGC	1484
Db	1440	CTATGCTGAGGCTATGAAGATCCCTACTACGACTATGATGATGCTATGACGTAAAGG	1499	
QY	1485	TAGAGGAAGAGGCTGTGAAGAGAGCAAGAGGAGTCTGCTCATCATCAAGAGTCGAGGGAGTGC	1544	
Db	1500	AGAGAGAGAGAGAGAGAGGGGGGGGGGAGAGTCTCTCAACCAAGAGGGGGGGAGACACC	1559	
QY	1545	TCCTCCCGCTGTGTAAGCCGCTTATTCACAGAGAGAGGCCCT--	CGATCAGCAAGAG	1601
Db	1560	ACCTCAAGAGGTAGAGCTGCTGCTATTCACAGAGGGGGGACCTTTGGAGCAACCAAGG	1619	
QY	1602	CGTTGCGGCTGTGAGAGAGGTCGCCAACAACAAGAGGCCCGCGGAGG	1649	
Db	1620	CTCAAGAGGTGGACAGAGGGGGTCTGTGCACACACAGAGAGAGGCCCGTGG	1667	
RESULT 15				
LOCUS	AC118904	151297 bp	DNA	linear
DEFINITION	Rattus norvegicus clone CH230-234U3, *** SEQUENCING IN PROGRESS			
ACCESSION	AC118904			
VERSION	AC118904.2	GI:21903156		
KEYWORDS	HTG, HTGS_PHASE1.			
SOURCE	Rattus norvegicus.			
ORGANISM	Rattus norvegicus			
	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			
	Rattus.			
REFERENCE	1 (bases 1 to 151297)			
AUTHORS	Munzly, D.M., Adams, C., Adlo-Oduola, B., Ali-osman, F.R., Allen, C., Albarino, S.L., Amarantunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbakats, J., Benton, J., Blagow, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Dem, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabliss, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Hollway, C., Hollins, B., Homel, F., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., Kling, L., Korrah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtenarge, O., Lieu, C., Liu, J., Liu, W., Louisedeg, H., Lozano, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Mareshwari, M., Mapua, P., Martin, R., Martinale, A., Martinez, E., Massey, E., McWhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miter, G., Miter, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neel, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkwo, S., Ogulu, M., Okunolu, G., Oregunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Plinius, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojoubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshtrati, N., Sison, I., Sodergren, B., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabot, P., Tameris, A., Tameris, K., Tang, H., Tatey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasequez, L., Verx, V., Villalón, D., Vinson, R., Wang, O., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleciyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.			
TITLE	Direct Submission			
JOURNAL	Unpublished			

REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	COMMENT
2 (bases 1 to 151297) Worley/K.C. Direct Submision Submitted (21-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 3 (bases 1 to 151297) Worley/K.C. Direct Submision Submitted (23-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA On Jul 18, 2002 this sequence version replaced gi:20258378.	Center: Baylor College of Medicine Genome Center Web site: <a href="http://www.hgsc.bcm.tmc.edu/">http://www.hgsc.bcm.tmc.edu/</a> Contact: <a href="mailto:hgsc-help@bcm.tmc.edu">hgsc-help@bcm.tmc.edu</a> Project Information Center project name: GWIS Center clone name: CH230-234J3 Summary Statistics Sequencing vector: Plasmid: Chemistry: Dye-terminator Big Dye: 100% of reads Assembly program: Phrap, version 0.990329 Consensus quality: 82728 bases at least Q40 Consensus quality: 89631 bases at least Q30 Consensus quality: 95421 bases at least Q20	
* NOTE: Estimated insert size may differ from sequence length * (see <a href="http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html">http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html</a> ) * NOTE: This is a 'working draft' sequence. It currently * consists of 81 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved.	1 1030 1029: contig of 1029 bp in length 1130 1129: gap of unknown length 2206 2205: contig of 1076 bp in length 2305 2305: gap of unknown length 2306 2305: contig of 1030 bp in length 3336 3335: gap of unknown length 3436 3435: contig of 1015 bp in length 4451 4450: gap of unknown length 4551 4550: contig of 1012 bp in length 5563 5562: gap of unknown length 6689 6688: contig of 1027 bp in length 6690 6689: gap of unknown length 7790 7789: contig of 1000 bp in length 7890 7889: gap of unknown length 7890 7889: contig of 1372 bp in length 9262 9261: gap of unknown length 9362 9361: gap of unknown length 10406 10406: contig of 1045 bp in length 10507 10506: gap of unknown length 10507 10513: contig of 1007 bp in length 11514 11513: gap of unknown length 11514 11513: gap of unknown length 12691 12691: contig of 1078 bp in length 12791 12791: gap of unknown length 12792 12791: contig of 1062 bp in length 13854 13853: gap of unknown length 13954 13953: gap of unknown length 15092 15092: contig of 1139 bp in length 15093 15092: gap of unknown length 15193 15193: contig of 1019 bp in length 16212 16211: gap of unknown length 16312 16311: gap of unknown length 17349 17349: contig of 1038 bp in length 17449 17449: gap of unknown length 18672 18672: contig of 1223 bp in length 18673 18672: gap of unknown length 18772 18772: gap of unknown length 20353 20353: contig of 1581 bp in length 20453 20453: gap of unknown length	

20454 21476: contig of 1023 bp in length  
21477 21576: gap of unknown length  
21577 22610: contig of 1034 bp in length  
22611 22710: gap of unknown length  
22711 23713: contig of 1003 bp in length  
23714 23813: gap of unknown length  
23814 25236: contig of 1423 bp in length  
25237 25336: gap of unknown length  
25337 26357: contig of 1021 bp in length  
26358 26457: gap of unknown length  
26458 27597: contig of 1140 bp in length  
27598 28843: contig of 1146 bp in length  
28844 28943: gap of unknown length  
28944 30133: contig of 1196 bp in length  
30140 30239: gap of unknown length  
30240 31240: contig of 1000 bp in length  
31240 31339: gap of unknown length  
31340 32369: contig of 1030 bp in length  
32370 32469: gap of unknown length  
32470 33484: contig of 1015 bp in length  
33485 33584: gap of unknown length  
33585 35005: contig of 1421 bp in length  
35006 36272: contig of 1166 bp in length  
36272 36371: gap of unknown length  
36372 37533: contig of 1162 bp in length  
37534 37633: gap of unknown length  
37634 38809: contig of 1176 bp in length  
38810 38909: gap of unknown length  
38910 40148: contig of 1239 bp in length  
40149 40248: gap of unknown length  
40249 41411: contig of 1163 bp in length  
41412 41512: gap of unknown length  
41512 43868: contig of 2357 bp in length  
43869 45611: contig of 1643 bp in length  
45612 45711: gap of unknown length  
45712 47108: contig of 1397 bp in length  
47109 47208: gap of unknown length  
47209 48273: contig of 1065 bp in length  
48274 48373: gap of unknown length  
48374 50006: contig of 1633 bp in length  
50007 50107: gap of unknown length  
50107 51371: contig of 1265 bp in length  
51372 51471: gap of unknown length  
51472 52752: contig of 1280 bp in length  
52753 52851: gap of unknown length  
52852 53943: contig of 1092 bp in length  
53944 54043: gap of unknown length  
54044 56187: contig of 2144 bp in length  
56188 56287: gap of unknown length  
56288 57740: contig of 1453 bp in length  
57741 57841: gap of unknown length  
57842 58957: contig of 1117 bp in length  
58958 59057: gap of unknown length  
59058 60557: contig of 1500 bp in length  
60559 60657: gap of unknown length  
60658 61991: contig of 1334 bp in length  
61992 62091: gap of unknown length  
62092 63961: contig of 1870 bp in length  
63962 64061: gap of unknown length  
64062 65338: contig of 1277 bp in length  
65339 65438: gap of unknown length  
65439 66681: contig of 1243 bp in length  
66682 66781: gap of unknown length  
66782 68027: contig of 1246 bp in length  
68028 68127: gap of unknown length  
68128 69486: contig of 1359 bp in length  
69487 69586: gap of unknown length  
69587 71278: contig of 1692 bp in length  
71279 71379: gap of unknown length  
71379 73777: contig of 2399 bp in length

Query Match 53.2%; Score 895.8; DB 2; Length 151297;  
Best Local Similarity 79.3%; Pred. No. 2.1e-193;  
Matches 1291; Conservative 0; Mismatches 279; Indels 59; Gaps 17;

59 CAGTTATCATTCAGAAAATTTTCAGACATGCTGTAGTGTGTTTACACAGAAAGTTG 118  
Db 43868 CACTATCCCATCCACAAATTTTTCAGACCT--TTGATGCTTAATATTCAAAAGTTG 43811

119 CTGAAAACATGATGAAATTTTACGTTGACAGGGCTAGTTGACATGATTTAGATGAAA 178  
Db 43810 CTGAAAACATGATGAAATTTTANG-TCGTGGCTAGTTGACATGATGATTTAGATGAAA 43752

179 GAGCTATCGAAGCTTTTAAAAGATTCATATGAAACGCGGATTTGGACGCTTCAACAGT 238  
Db 43751 AAGCTATCAAAAGCTTAAAAGATTCATATGAAACATGATGATGAAAGCTTCAACATT 43692

239 TTAAGACAGTGAATCTCTCATGTTTCAGAAACAAAGTGCTTTTATGAGATGATGA 298  
Db 43691 TTAAGACAGTGAATCTCTCTCTG-----CCCTTTTATGAGATGATGA 43646

299 AGACTTACAGCAG 358  
Db 43645 AGACTTACAGCAG 43587

359 ATGAGGCAAGATTTAAGGACCTTTTGGAAAAGACAGGCTACACATTTGATGATCAG 418  
Db 43586 ATGAGGCAAGATTTAAGGACCTTTTGGAAAAGAGAGAGAGAGAGAGAGAGAGAG 43527

419 GTGAGAGAGATTTGAG 477  
Db 43526 GACAGAGAGATTTGAG 43467

478 GTTGCACTGAGATTTTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 537  
Db 43466 GTTGCACTGAGATTTTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 43407

538 CCAATATTTAG 597  
Db 43406 CCAATATTTAG 43347

598 GGTCTCAAG 657  
Db 43346 GGTCTCAAG 43290

658 GTTAACTGTAT 717  
Db 43289 GTTAACTGT-----ATCATGAAATTCATTTGGAAGAGAGAGAGAGAGAGAGAG 43236

718 GTTGCCCAAGATTTAGGCTTTTGTGGGCTGATTCCTTAAAGATTTAAACCAAGAGAGAT 777  
Db 43235 GTTGCCCAAGATTTAGG--GGGGGGGGGCTCGTTTCAAGATTTAAACCAAGAGATTT 43178

778 CTTGAGGAAATTTAGCAAG 837  
Db 43177 CTTGAGGAAATTTAGCAAG 43118

838 GATGACAAAG 897  
Db 43117 GATGACAAAG 43059

898 GCCCAG 957  
Db 43058 GCCCAG 42999

958 GTTGAAG 1017  
Db- 42998 GTTGAAG 42945

1018 CTGTTTGTAGCAACCTTGGCAACAGGTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1077  
Db 42944 CTGTTTGTAGCAACCTTGGCAACATATATATATATATATATATATATATATATATAT 42885





**THIS PAGE BLANK (USPTO)**